02

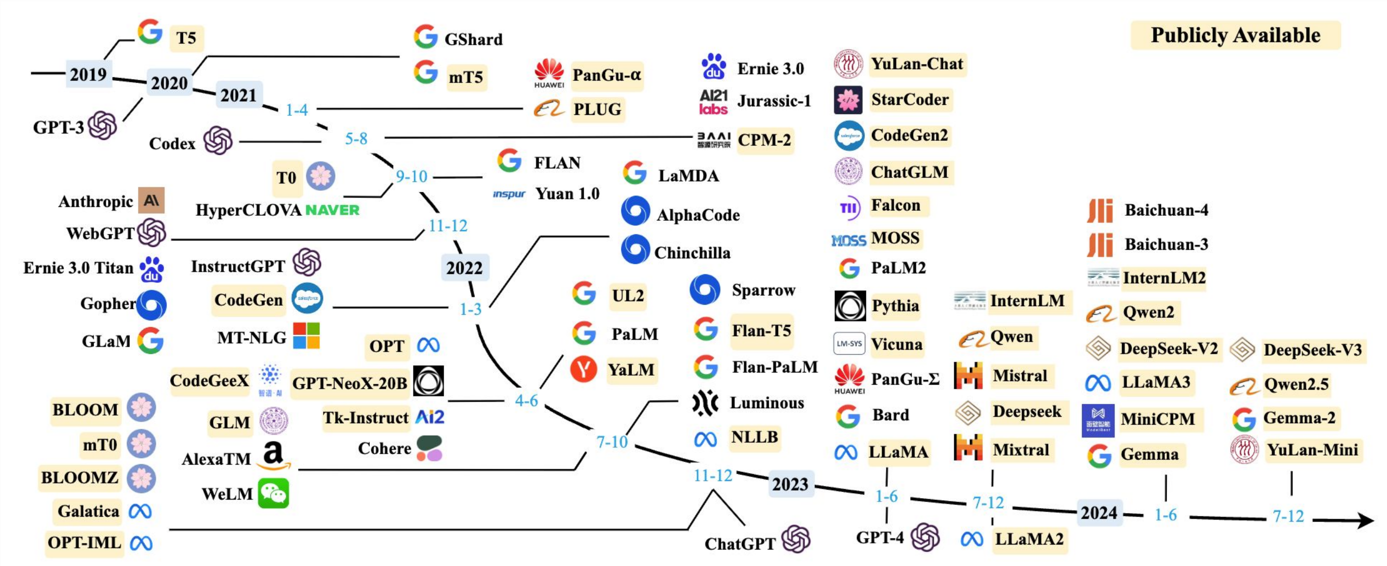
LLM

-based Generative Recommendation

The Rise of Large Language Models

Transformer

2017



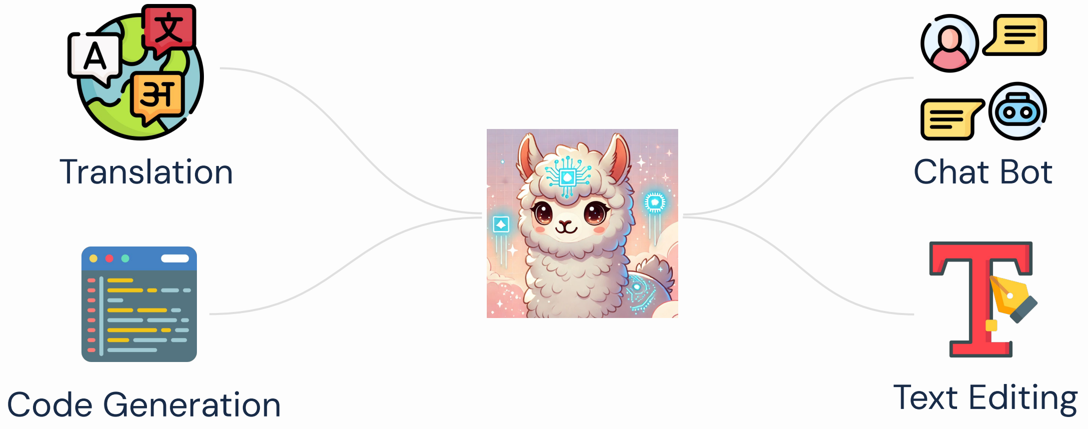
O3, R1…

2025

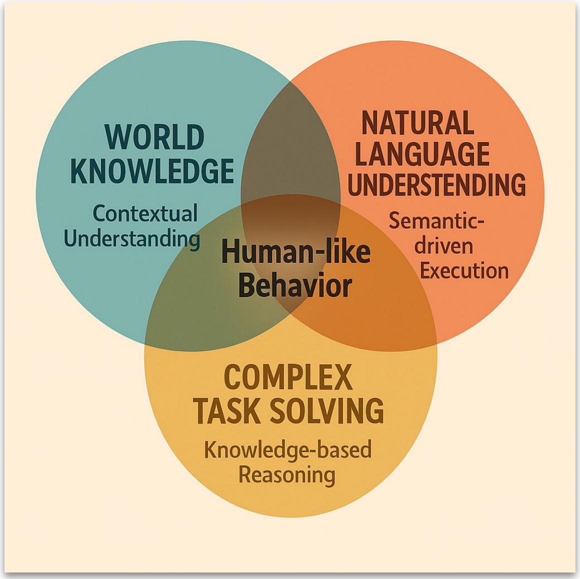
LLMs are developing so fast recently…

Large Language Models

LLMs are machine learning models that can perform a variety of natural language processing (NLP) tasks



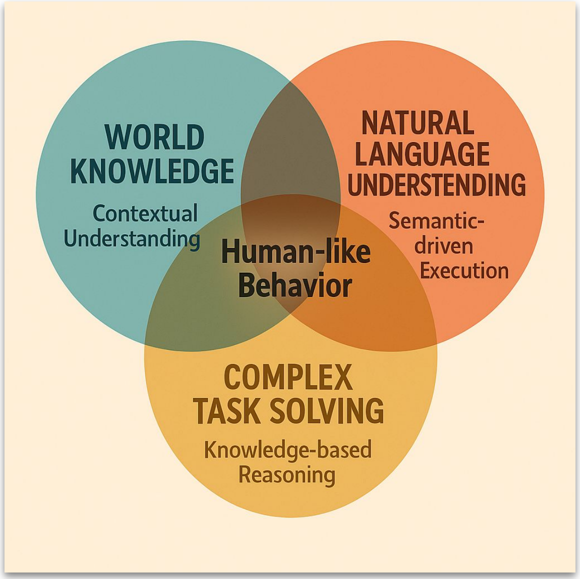
Large Language Models



Key features of LLMs:

- World knowledge.
- Natural language understanding.
- Human-like behavior.

Large Language Models



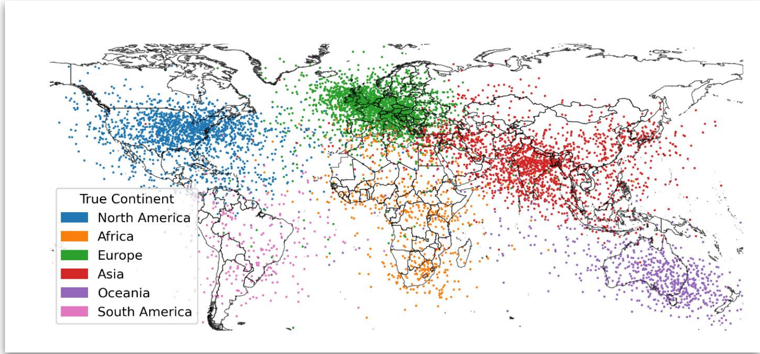
Key features of LLMs:

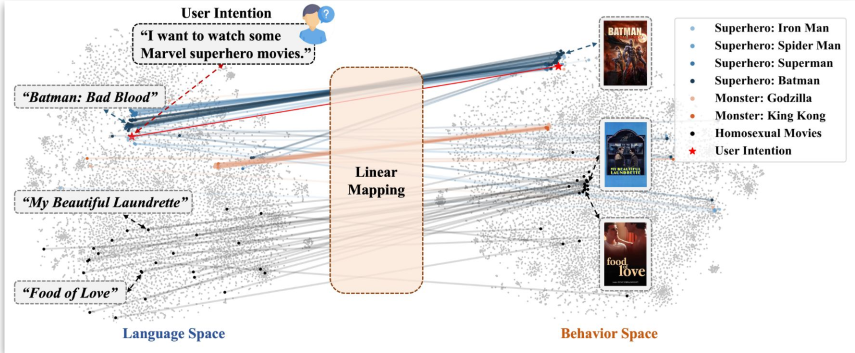
- World knowledge.
- Natural language understanding.
- Human-like behavior.

How can these features benefit recommender systems?

Benefits of LLMs for Recommendation

(1) World knowledge – from pretraining





In space

In recommendation

Gurnee et al. Language Models Represent Space and Time. ICLR 2024.
Sheng et al. Language Representations Can be What Recommenders Need: Findings and Potentials. ICLR 2025.

Benefits of LLMs for Recommendation

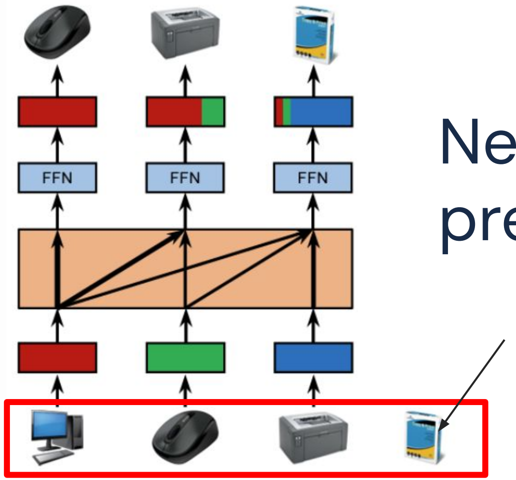
(1) World knowledge

LLM as sequential recommender

-> Alleviating the data sparsity of ID-based interactions in recommendation

Benefits of LLMs for Recommendation

(1) World knowledge



Next ID prediction

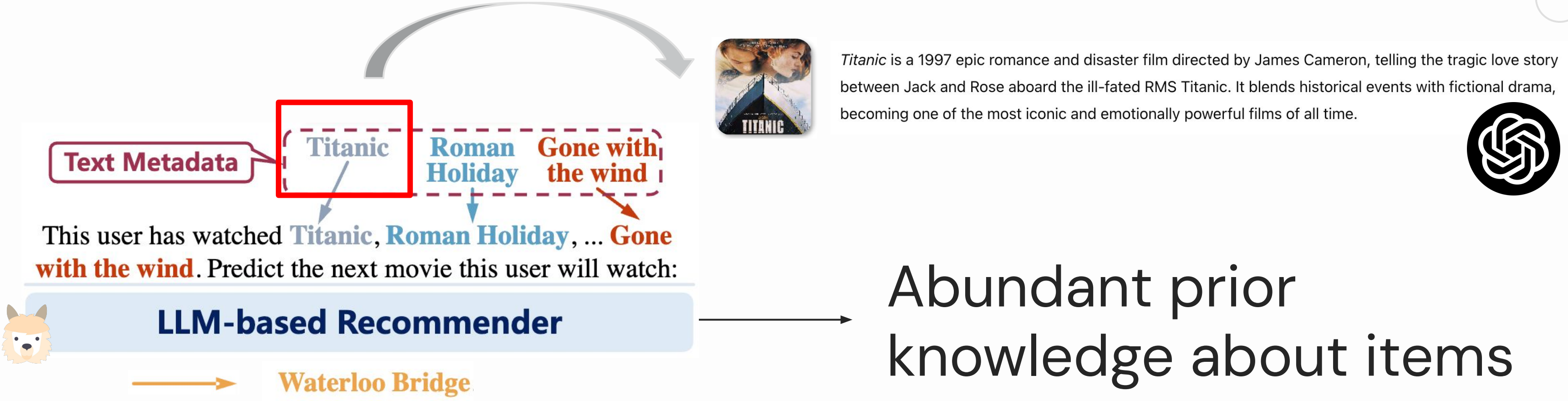
Item IDs

ID-based item modeling lack semantic meanings

Example: SASRec [*ICDM'18*]

Benefits of LLMs for Recommendation

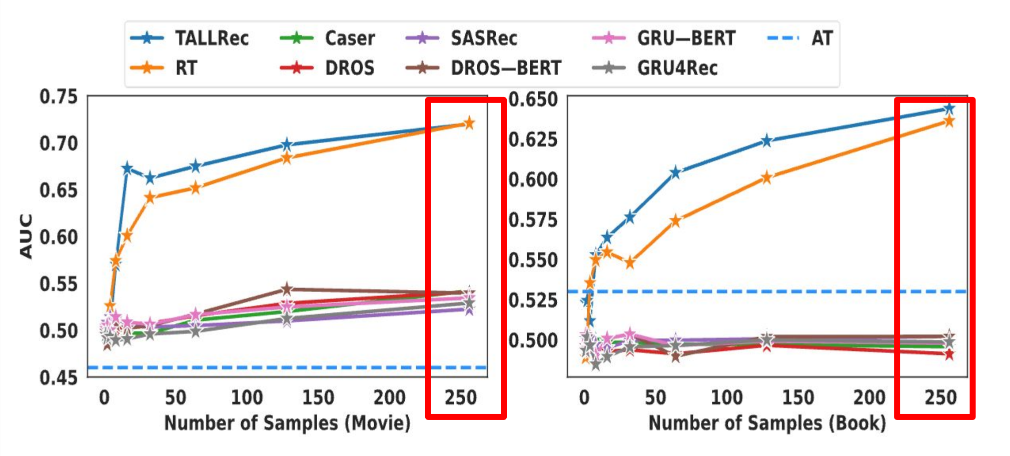
(1) World knowledge



Abundant prior knowledge about items

Benefits of LLMs for Recommendation

(1) World knowledge



Few data -> a good recommender

Benefits of LLMs for Recommendation

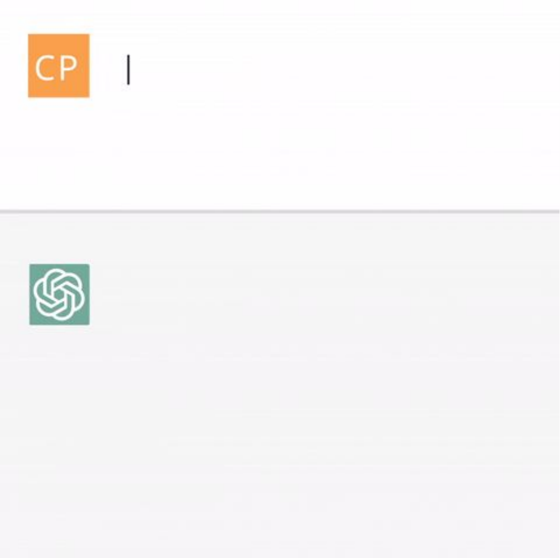
(1) World knowledge

LLM as sequential recommender

Lower data requirement
Cross-domain ability
Cold-start ability
...

Benefits of LLMs for Recommendation

(2) Natural language understanding & generation



LLMs can interact with users fluently

Benefits of LLMs for Recommendation

(2) Natural language understanding & generation

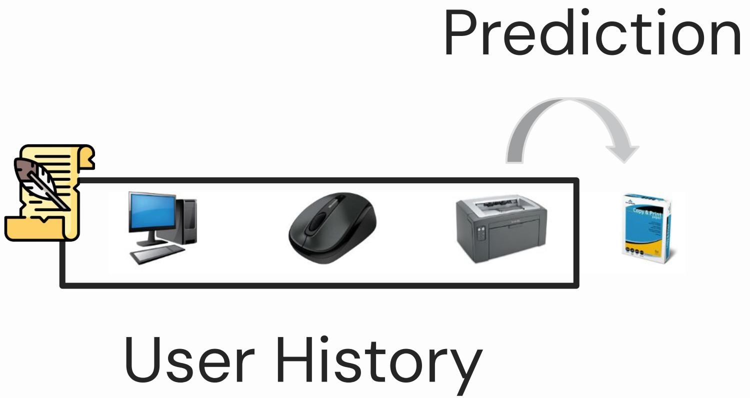
LLM as conversational recommender

-> Towards more interactive recommender systems

Benefits of LLMs for Recommendation

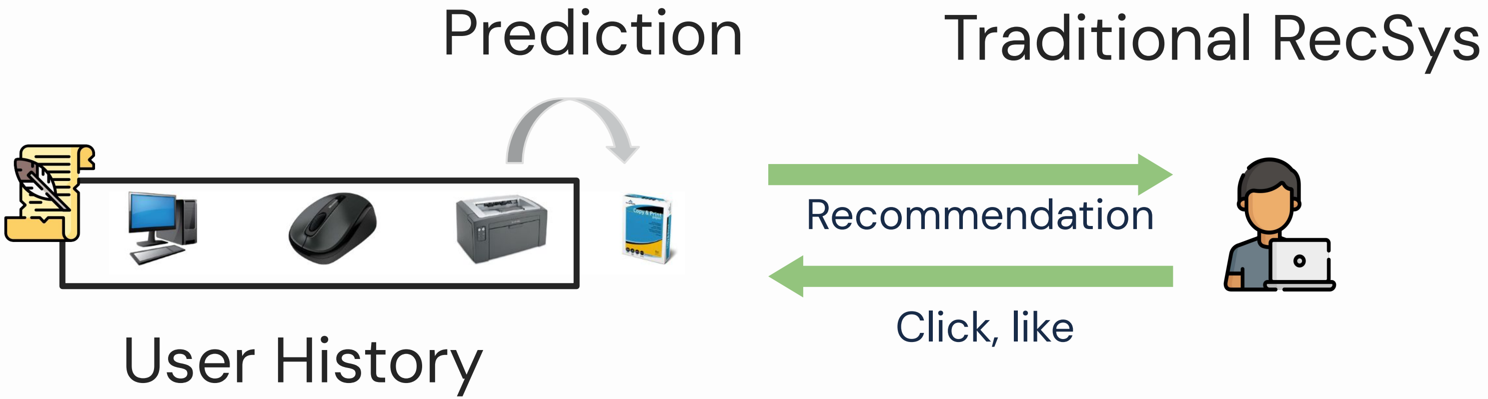
(2) Natural language understanding & generation

Traditional RecSys



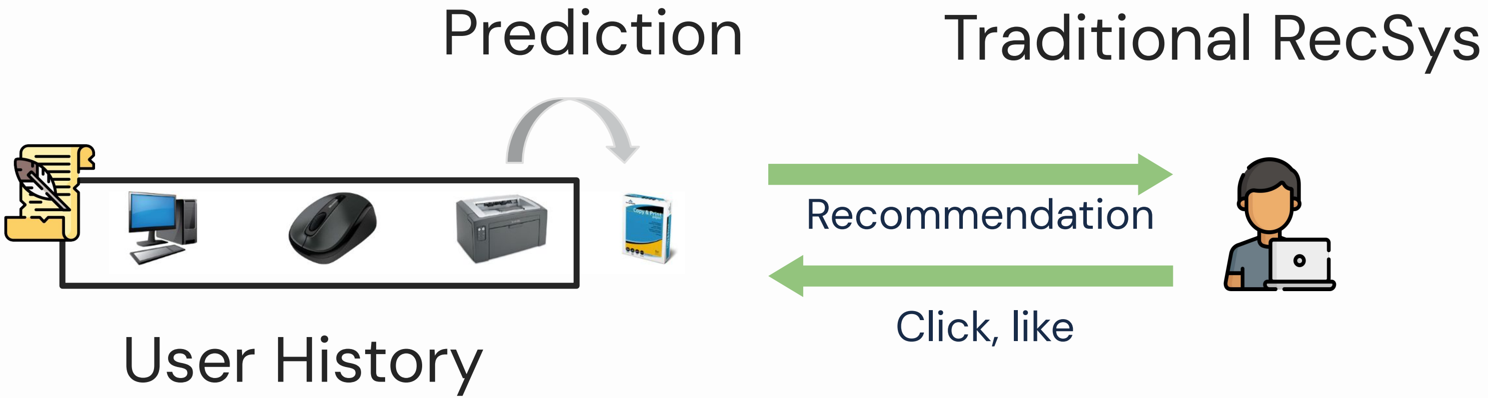
Benefits of LLMs for Recommendation

(2) Natural language understanding & generation



Benefits of LLMs for Recommendation

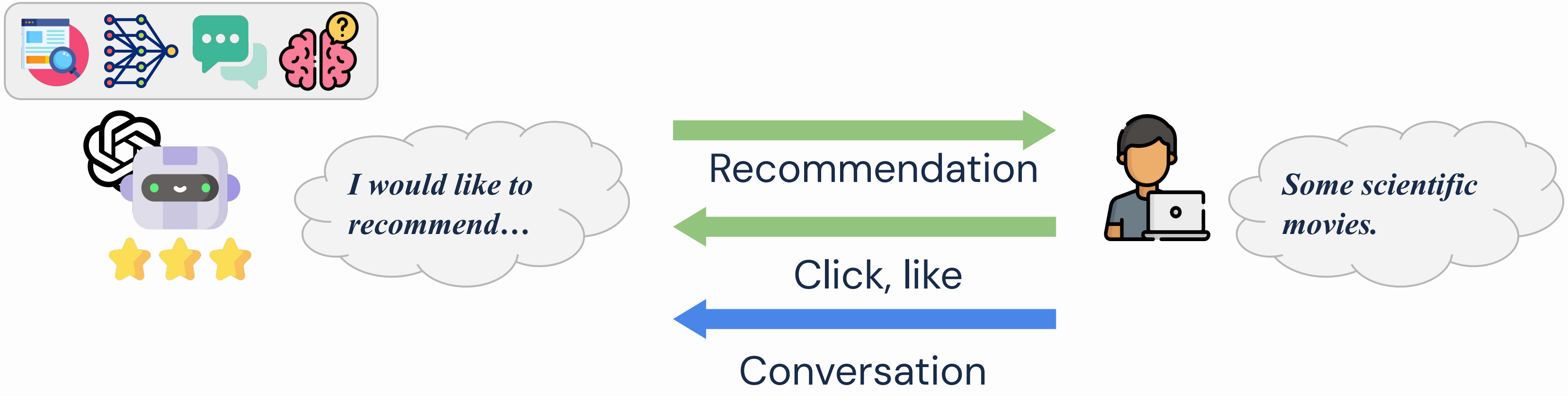
(2) Natural language understanding & generation



Passive recommendation!

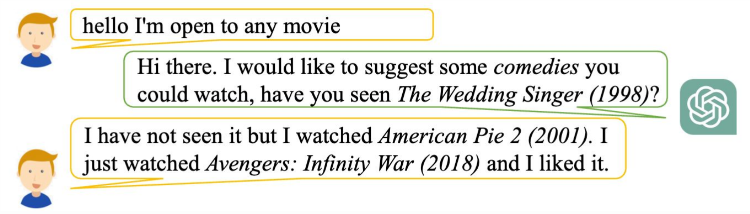
Benefits of LLMs for Recommendation

(2) Natural language understanding & generation



Benefits of LLMs for Recommendation

(2) Natural language understanding & generation



LLM as conversational recommender

Interactive
User-friendly
More accurate
...

Benefits of LLMs for Recommendation

(3) Human-like behavior

Park et al. Generative Agents: Interactive Simulacra of Human Behavior. UIST 2023

Benefits of LLMs for Recommendation

(3) Human-like behavior

Generative Agents can (mostly) simulate human behaviors

- Cooperation
- Organization

Park et al. Generative Agents: Interactive Simulacra of Human Behavior. UIST 2023

Benefits of LLMs for Recommendation

(3) Human-like behavior

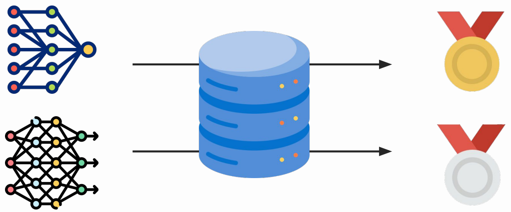
LLM as user simulator

-> Simulating user behaviors for evaluating recommenders.

Benefits of LLMs for Recommendation

(3) Human-like behavior

Offline recommender evaluation

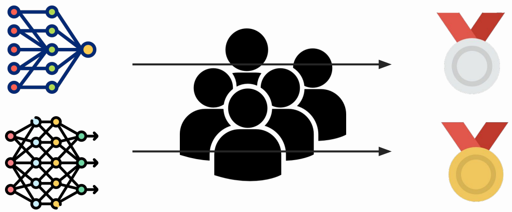


Inaccurate, but affordable

Benefits of LLMs for Recommendation

(3) Human-like behavior

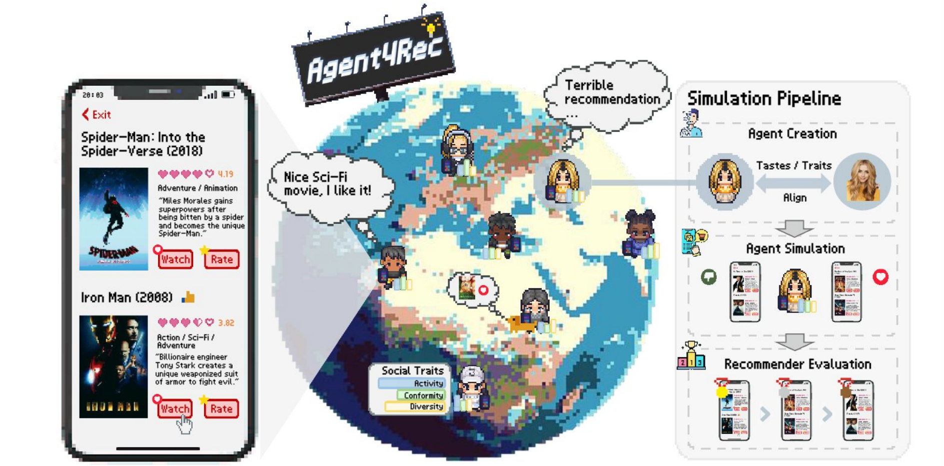
Online recommender evaluation



Accurate, but
costly

Benefits of LLMs for Recommendation

(3) Human-like behavior



LLM as user simulator

Faithful
Affordable
Controllable
...

Part 1: LLM as Sequential Recommender

(i) **Early efforts**: Pretrained LLMs for recommendation;

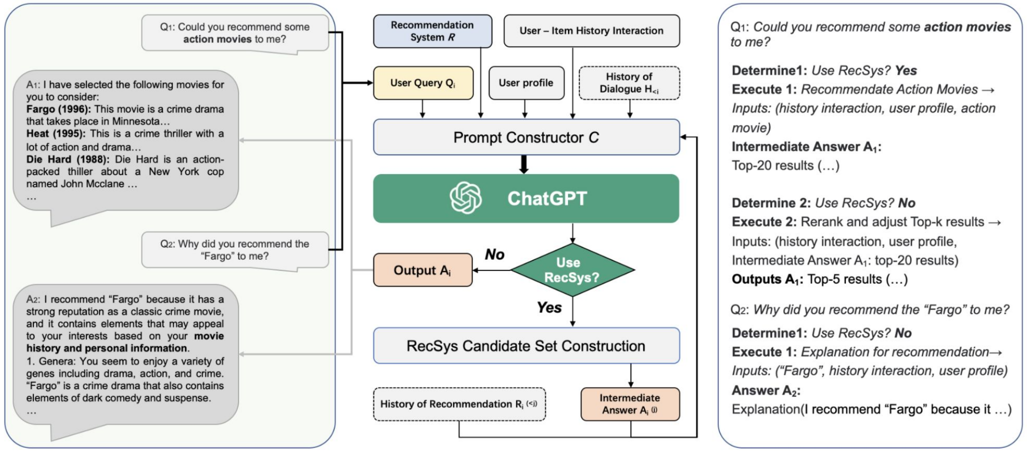
Early efforts

- Directly use freezed LLMs (e.g., GPT 4) for recommendation.

Early efforts

Prompt Engineering + In-Context Learning (ChatRec)

Key idea: LLMs as the recsys controller

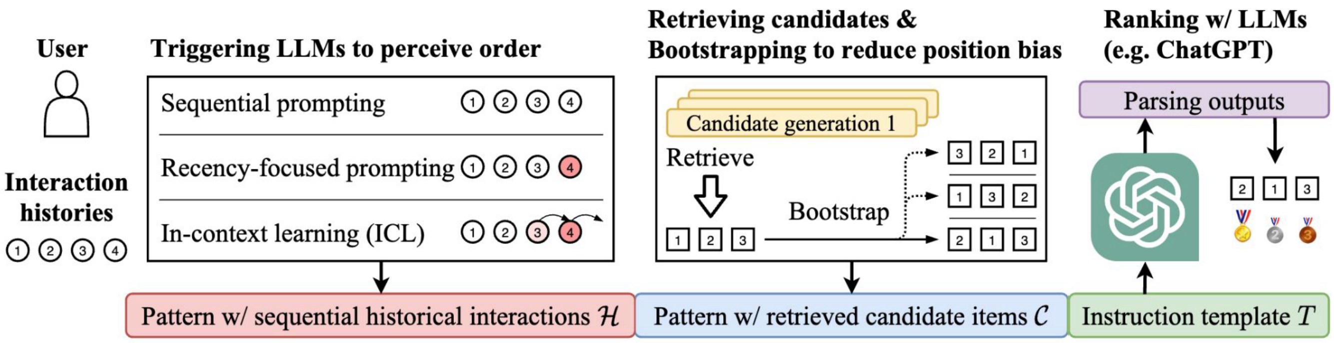


Gao et al. Chat-REC: Towards Interactive and Explainable LLMs-Augmented Recommender System. arXiv 2303.14524.

Early efforts

Prompt Engineering + In-Context Learning (LLMRank)

Key idea: LLMs as the reranker



Early efforts

- Directly use freezed LLMs (e.g., GPT 4) for recommendation.
- A performance gap compared to traditional recommenders exists.

Early efforts

Sub-optimal performance comparing to SASRec!

Performance of LLMRank

	Method	ML-1M N@1	ML-1M N@5	ML-1M N@10	ML-1M N@20	Games N@1	Games N@5	Games N@10	Games N@20
full	Pop	0.08	1.20	4.13	5.79	0.13	1.00	2.27	2.62
full	BPRMF [49]	0.26	1.69	4.41	6.04	0.55	1.98	**2.96**	**3.19**
full	SASRec [33]	**3.76**	**9.79**	**10.45**	**10.56**	**1.33**	**3.55**	**4.02**	**4.11**
zero-shot	BM25 [50]	0.26	0.87	2.32	5.28	0.18	1.07	1.80	2.55
zero-shot	UniSRec [30]	0.88	3.46	5.30	6.92	0.00	1.86	2.03	2.31
zero-shot	VQ-Rec [29]	0.20	1.60	3.29	5.73	0.20	1.21	1.91	2.64
zero-shot	Ours	**1.74**	**5.22**	**6.91**	**7.90**	**0.90**	**2.26**	2.80	3.08

Hou et al. Large Language Models are Zero-Shot Rankers for Recommender Systems. ECIR 2024.

Early efforts

Sub-optimal performance comparing to SASRec!

Aligning LLMs for recommendation tasks is necessary!

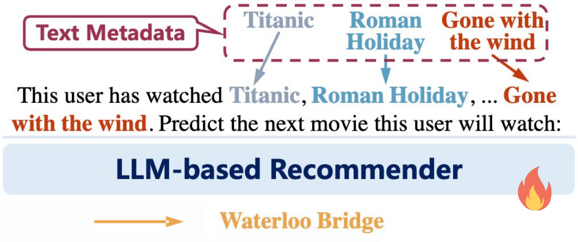
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
zero-shot	BM25 [50]	0.26	0.87	2.32	5.28	0.18	1.07	1.80	2.55
	UniSRec [30]	0.88	3.46	5.30	6.92	0.00	1.86	2.03	2.31
	VQ-Rec [29]	0.20	1.60	3.29	5.73	0.20	1.21	1.91	2.64
	Ours	**1.74**	**5.22**	**6.91**	**7.90**	**0.90**	**2.26**	2.80	3.08

Part 1: LLM as Sequential Recommender

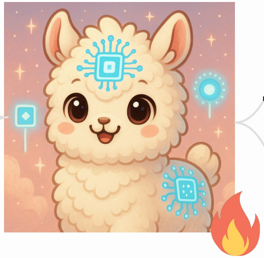
(i) Early efforts: Pretrained LLMs for recommendation;

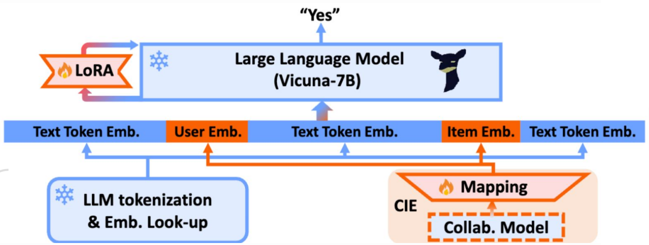
(ii) **Aligning** LLMs for recommendation;

Aligning LLMs for recommendation

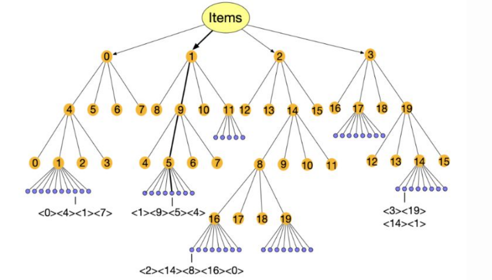


Pure text-based

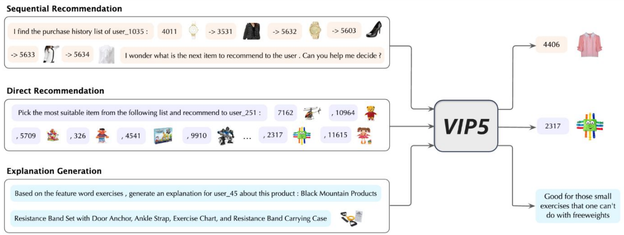




+ Collaborative embeddings

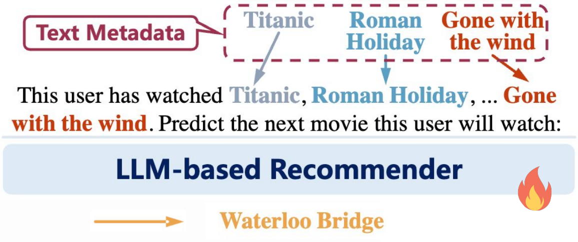


+ External item tokens

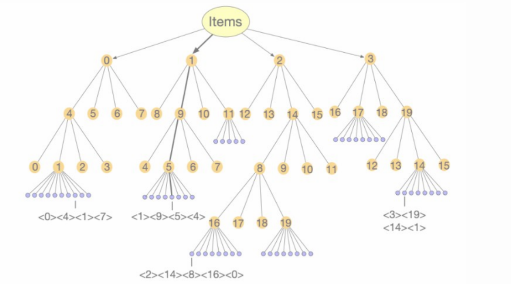


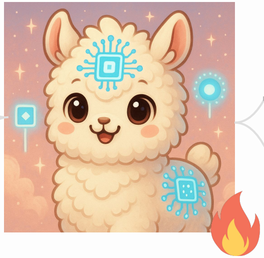
+ Multimodal information

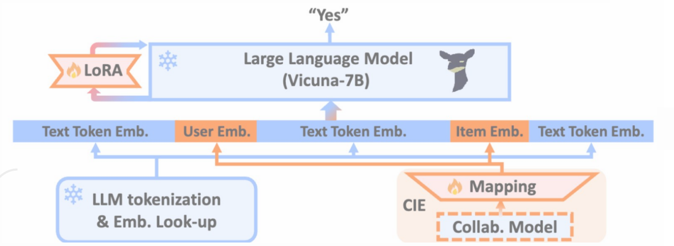
Aligning LLMs for recommendation



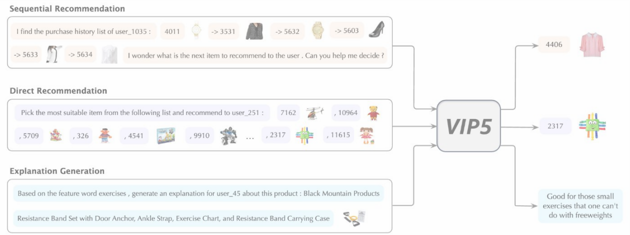
Pure text-based







+ Collaborative embeddings

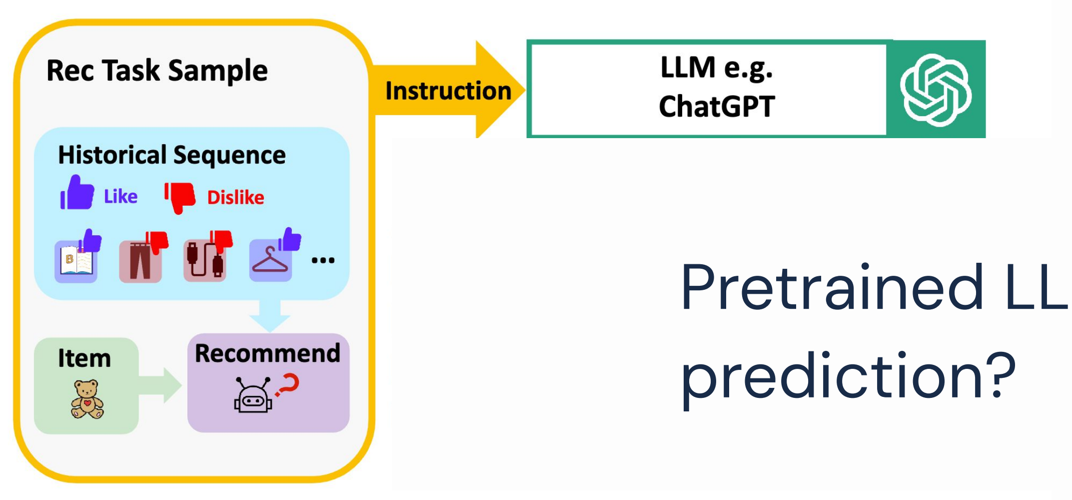


+ External item tokens

+ Multimodal information

Aligning LLMs for recommendation

(1) Pure text-based (TALLRec)

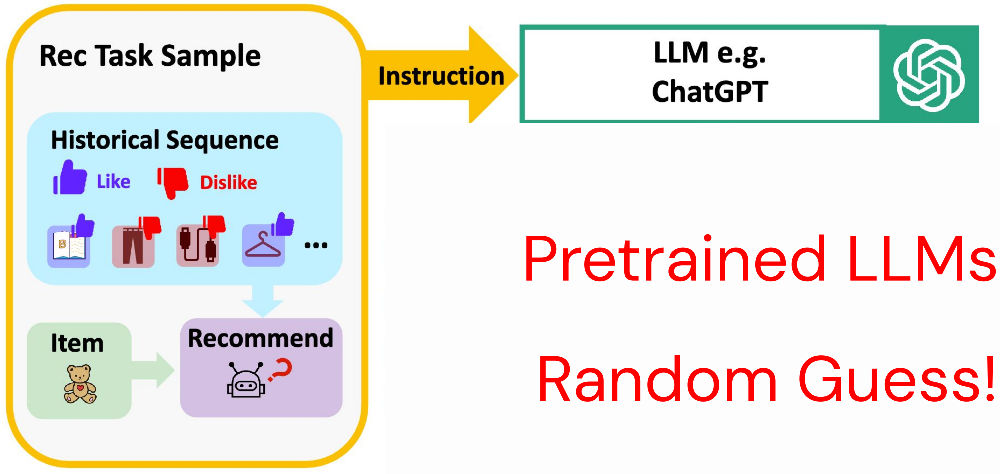


Pretrained LLMs for CTR prediction?

Bao et al. TALLRec: An Effective and Efficient Tuning Framework to Align Large Language Model with Recommendation. RecSys 2023.

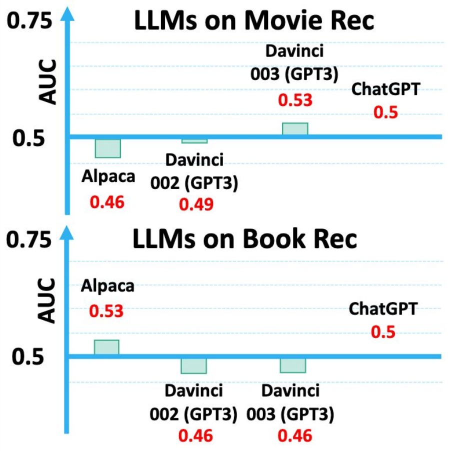
Aligning LLMs for recommendation

(1) Pure text-based (TALLRec)



Pretrained LLMs:

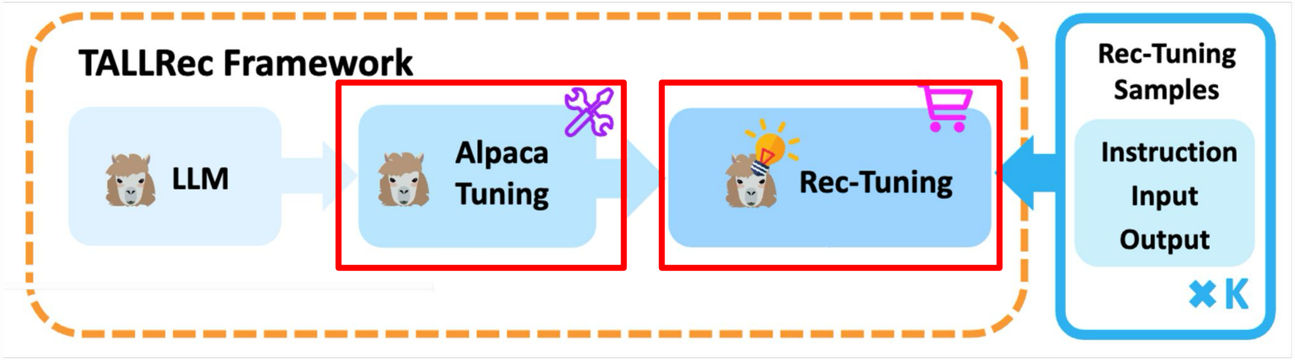
Random Guess!



Bao et al. TALLRec: An Effective and Efficient Tuning Framework to Align Large Language Model with Recommendation. RecSys 2023.

Aligning LLMs for recommendation

(1) Pure text-based (TALLRec)

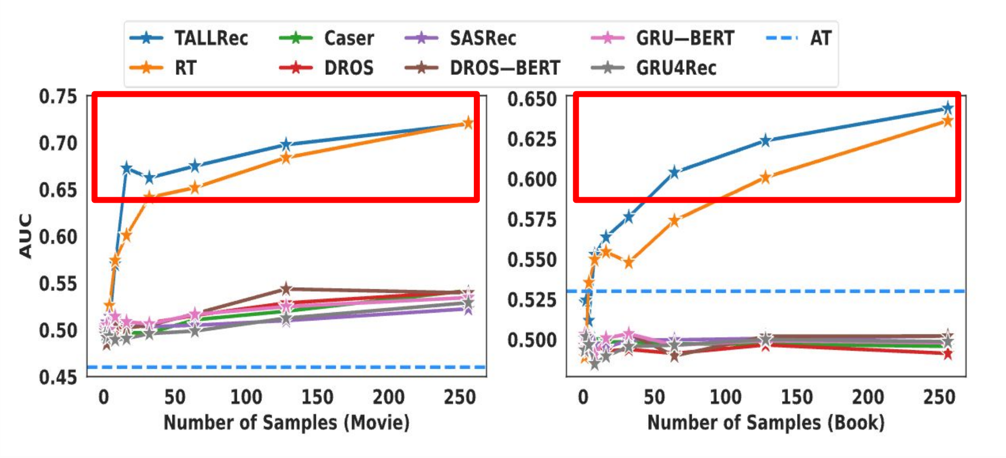


General task alignment -> Recommendation alignment

Bao et al. TALLRec: An Effective and Efficient Tuning Framework to Align Large Language Model with Recommendation. RecSys 2023.

Aligning LLMs for recommendation

(1) Pure text-based (TALLRec)

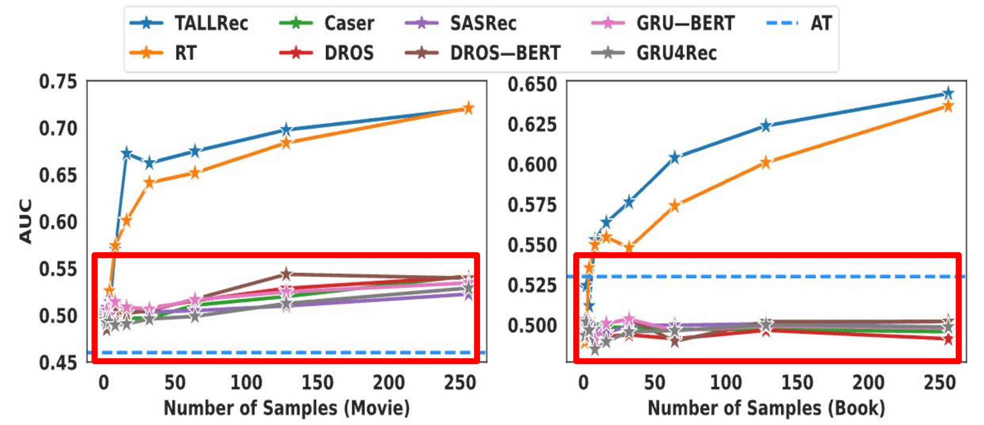


Few training data -> Huge improvements

Bao et al. TALLRec: An Effective and Efficient Tuning Framework to Align Large Language Model with Recommendation. RecSys 2023.

Aligning LLMs for recommendation

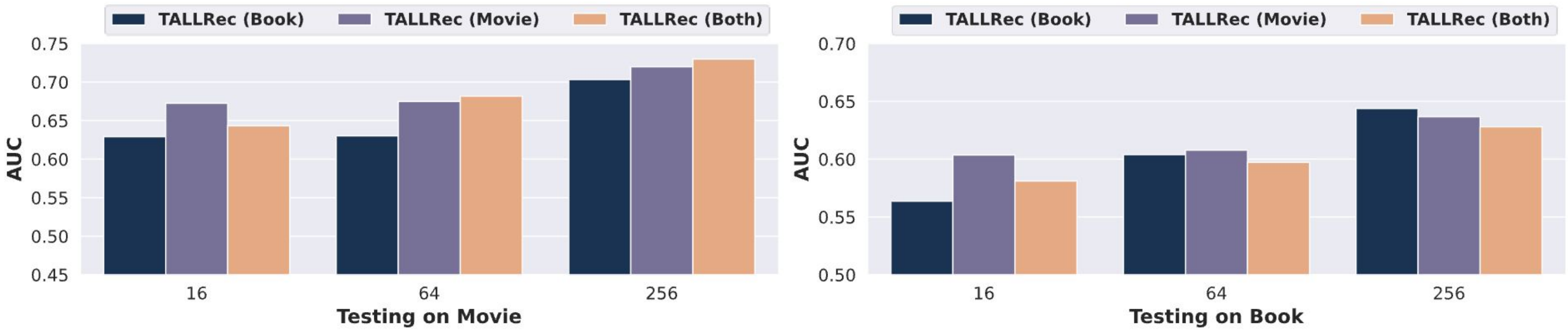
(1) Pure text-based (TALLRec)



Traditional recommenders: suffer from too-sparse supervision signals

Aligning LLMs for recommendation

(1) Pure text-based (TALLRec)

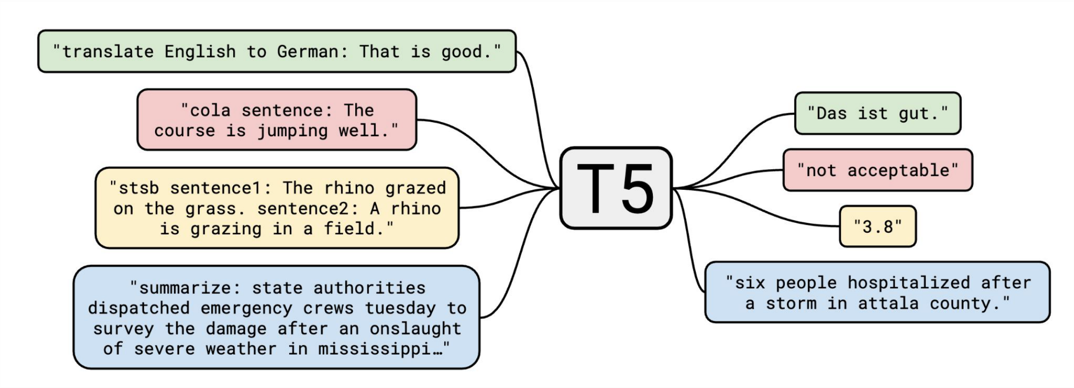


Cross-domain generalization

Bao et al. TALLRec: An Effective and Efficient Tuning Framework to Align Large Language Model with Recommendation. RecSys 2023.

Aligning LLMs for recommendation

(1) Pure text-based – Multiple rec taks

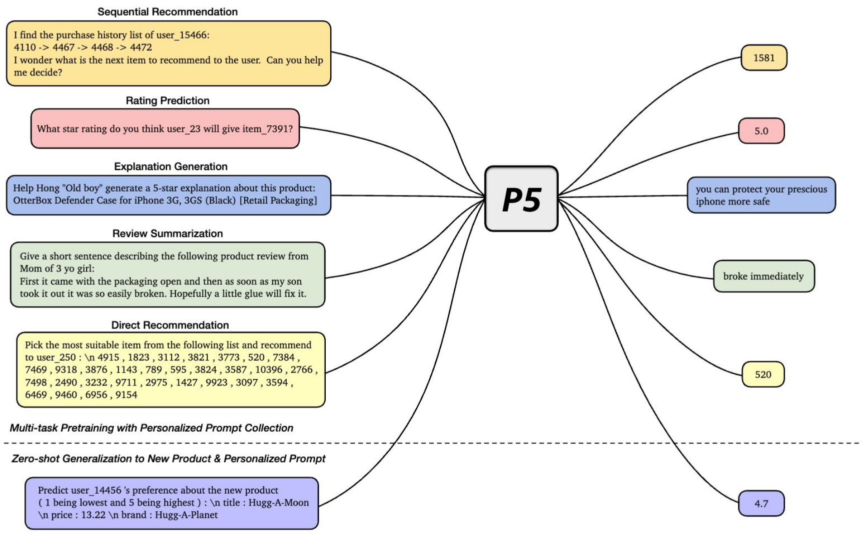


Unified language modeling in NLP

Raffel et al. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. JMLR 2020.

Aligning LLMs for recommendation

(1) Pure text-based – Multiple rec taks

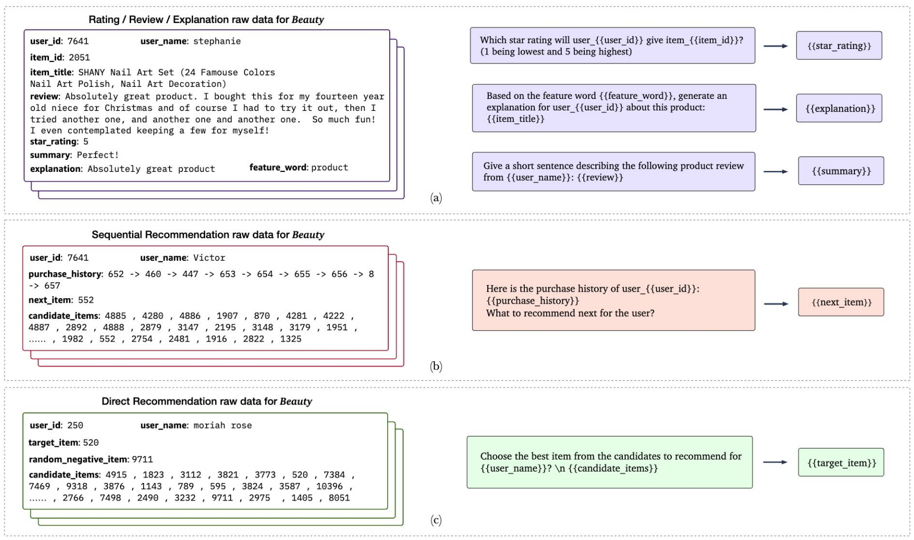


Multi-task alignment (P5)

-> general recommender

Aligning LLMs for recommendation

(1) Pure text-based – Multiple rec taks



Training on different task prompts -> multiple recommendation abilities.

Aligning LLMs for recommendation

(1) Pure text-based – Multiple rec taks

Table 6: Performance comparison on review summarization (%).

Methods	Sports				Beauty				Toys			
	BLUE2	ROUGE1	ROUGE2	ROUGEL	BLUE2	ROUGE1	ROUGE2	ROUGEL	BLUE2	ROUGE1	ROUGE2	ROUGEL
T0 (4-1)	2.1581	2.2695	0.5694	1.6221	1.2871	1.2750	0.3904	0.9592	2.2296	2.4671	0.6482	1.8424
GPT-2 (4-1)	0.7779	4.4534	1.0033	1.9236	0.5879	3.3844	0.6756	1.3956	0.6221	3.7149	0.6629	1.4813
P5-S (4-1)	2.4962	11.6701	2.7187	10.4819	**2.1225**	**8.4205**	**1.6676**	**7.5476**	**2.4752**	**9.4200**	**1.5975**	**8.2618**
P5-B (4-1)	**2.6910**	**12.0314**	**3.2921**	**10.7274**	1.9325	8.2909	1.4321	7.4000	1.7833	8.7222	1.3210	7.6134

Table 7: Performance comparison on direct recommendation.

Methods	Sports					Beauty					Toys				
	HR@1	HR@5	NDCG@5	HR@10	NDCG@10	HR@1	HR@5	NDCG@5	HR@10	NDCG@10	HR@1	HR@5	NDCG@5	HR@10	NDCG@10
BPR-MF	0.0314	0.1404	0.0848	0.2563	0.1220	0.0311	0.1426	0.0857	0.2573	0.1224	0.0233	0.1066	0.0641	0.2003	0.0940
BPR-MLP	0.0351	0.1520	0.0927	0.2671	0.1296	0.0317	0.1392	0.0848	0.2542	0.1215	0.0252	0.1142	0.0688	0.2077	0.0988
SimpleX	0.0331	**0.2362**	**0.1505**	0.3290	0.1800	0.0325	0.2247	0.1441	0.3090	0.1711	0.0268	**0.1958**	**0.1244**	**0.2662**	**0.1469**
P5-S (5-1)	0.0638	0.2096	0.1375	0.3143	0.1711	0.0600	0.2021	0.1316	0.3121	0.1670	0.0405	0.1538	0.0969	0.2405	0.1248
P5-B (5-1)	0.0245	0.0816	0.0529	0.1384	0.0711	0.0224	0.0904	0.0559	0.1593	0.0780	0.0187	0.0827	0.0500	0.1543	0.0729
P5-S (5-4)	0.0701	0.2241	0.1483	**0.3313**	**0.1827**	**0.0862**	**0.2448**	**0.1673**	**0.3441**	**0.1993**	0.0413	0.1411	0.0916	0.2227	0.1178
P5-B (5-4)	0.0299	0.1026	0.0665	0.1708	0.0883	0.0506	0.1557	0.1033	0.2350	0.1287	0.0435	0.1316	0.0882	0.2000	0.1102
P5-S (5-5)	0.0574	0.1503	0.1050	0.2207	0.1276	0.0601	0.1611	0.1117	0.2370	0.1360	0.0440	0.1282	0.0865	0.2011	0.1098
P5-B (5-5)	0.0641	0.1794	0.1229	0.2598	0.1488	0.0588	0.1573	0.1089	0.2325	0.1330	0.0386	0.1122	0.0756	0.1807	0.0975
P5-S (5-8)	0.0567	0.1514	0.1049	0.2196	0.1269	0.0571	0.1566	0.1078	0.2317	0.1318	**0.0451**	0.1322	0.0889	0.2023	0.1114
P5-B (5-8)	**0.0726**	0.1955	0.1355	0.2802	0.1627	0.0608	0.1564	0.1096	0.2300	0.1332	0.0389	0.1147	0.0767	0.1863	0.0997

Table 3: Performance comparison on sequential recommendation.

Methods	Sports				Beauty				Toys			
	HR@5	NDCG@5	HR@10	NDCG@10	HR@5	NDCG@5	HR@10	NDCG@10	HR@5	NDCG@5	HR@10	NDCG@10
Caser	0.0116	0.0072	0.0194	0.0097	0.0205	0.0131	0.0347	0.0176	0.0166	0.0107	0.0270	0.0141
HGN	0.0189	0.0120	0.0313	0.0159	0.0325	0.0206	0.0512	0.0266	0.0321	0.0221	0.0497	0.0277
GRU4Rec	0.0129	0.0086	0.0204	0.0110	0.0164	0.0099	0.0283	0.0137	0.0097	0.0059	0.0176	0.0084
BERT4Rec	0.0115	0.0075	0.0191	0.0099	0.0203	0.0124	0.0347	0.0170	0.0116	0.0071	0.0203	0.0099
FDSA	0.0182	0.0122	0.0288	0.0156	0.0267	0.0163	0.0407	0.0208	0.0228	0.0140	0.0381	0.0189
SASRec	0.0233	0.0154	0.0350	0.0192	0.0387	0.0249	0.0605	0.0318	0.0463	0.0306	0.0675	0.0374
S^3-Rec	0.0251	0.0161	0.0385	0.0204	0.0387	0.0244	0.0647	0.0327	0.0443	0.0294	0.0700	0.0376
P5-S (2-3)	0.0272	0.0169	0.0361	0.0198	0.0503	0.0370	0.0659	0.0421	**0.0648**	**0.0567**	**0.0709**	**0.0587**
P5-B (2-3)	0.0364	0.0296	0.0431	0.0318	**0.0508**	**0.0379**	**0.0664**	**0.0429**	0.0608	0.0507	0.0688	0.0534
P5-S (2-13)	0.0258	0.0159	0.0346	0.0188	0.0490	0.0358	0.0646	0.0409	0.0647	0.0566	0.0705	0.0585
P5-B (2-13)	**0.0387**	**0.0312**	**0.0460**	**0.0336**	0.0493	0.0367	0.0645	0.0416	0.0587	0.0486	0.0675	0.0536

Table 4: Performance comparison on explanation generation (%).

Methods	Sports				Beauty				Toys			
	BLUE4	ROUGE1	ROUGE2	ROUGEL	BLUE4	ROUGE1	ROUGE2	ROUGEL	BLUE4	ROUGE1	ROUGE2	ROUGEL
Attn2Seq	0.5305	12.2800	1.2107	9.1312	0.7889	12.6590	1.6820	9.7481	1.6238	13.2245	2.9942	10.7398
NRT	0.4793	11.0723	1.1304	7.6674	0.8295	12.7815	1.8543	9.9477	1.9084	13.5231	3.6708	11.1867
PETER	0.7112	12.8944	1.3283	9.8635	1.1541	14.8497	2.1413	11.4143	1.9861	14.2716	3.6718	11.7010
P5-S (3-3)	**1.0447**	**14.9048**	**2.1297**	**11.1778**	**1.2237**	**17.6938**	**2.2489**	**12.8606**	2.2892	**15.4505**	3.6974	**12.1718**
P5-B (3-3)	1.0407	14.1589	2.1220	10.6096	0.9742	16.4530	1.8858	11.8765	**2.3185**	15.3474	**3.7209**	12.1312
PETER+	**2.4627**	**24.1181**	5.1937	**18.4105**	**3.2606**	25.5541	5.9668	19.7168	**4.7919**	28.3083	9.4520	**22.7017**
P5-S (3-9)	1.4101	23.5619	**5.4196**	17.6245	1.9788	**25.6253**	**6.3678**	**19.9497**	4.1222	**28.4088**	9.5432	22.6064
P5-B (3-9)	1.4689	23.5476	5.3926	17.5852	1.8765	25.1183	6.0764	19.4488	3.8933	27.9916	9.5896	22.2178
P5-S (3-12)	1.3212	23.2474	5.3461	17.3780	1.9425	25.1474	6.0551	19.5601	4.2764	28.1897	9.1327	22.2514
P5-B (3-12)	1.4303	23.3810	5.3239	17.4913	1.9031	25.1763	6.1980	19.5188	3.5861	28.1369	**9.7562**	22.3056

Single LLM -> Effective on various recommendation tasks

Geng et al. Recommendation as Language Processing (RLP): A Unified Pretrain, Personalized Prompt & Predict Paradigm (P5). RecSys 2022.

Aligning LLMs for recommendation

(1) Pure text-based (P5)

Multi-scenario Recommendation: The items the user has recently clicked on are as follows: {USER BEHAVIOR SEQUENCE}. In scenario {SCENE}, please recommend items.

Multi-objective Recommendation: The items the user has recently clicked on are as follows: {USER BEHAVIOR SEQUENCE}. Please find items that the user will {ACTION}.

Long-tail Item Recommendation: The items the user has recently clicked on are as follows: {USER BEHAVIOR SEQUENCE}. Please recommend long-tail items.

Serendipity Recommendation: The items the user has recently clicked on are as follows: {USER BEHAVIOR SEQUENCE}. Please recommend some new item categories.

Long-term Recommendation: The items the user has recently clicked on are as follows: {USER BEHAVIOR SEQUENCE}. Please find items that match the user's long-term interests.

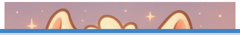
Search Problem: The items the user has recently clicked on are as follows: {USER BEHAVIOR SEQUENCE}. Please recommend items that match {QUERY}.

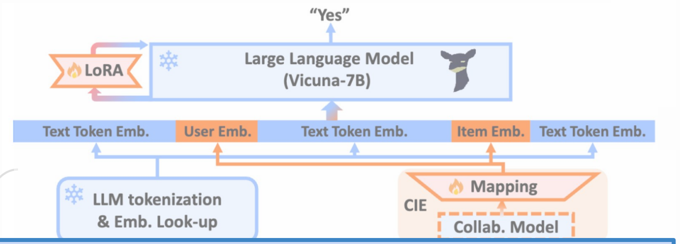
URM:

Unify recommendation & search

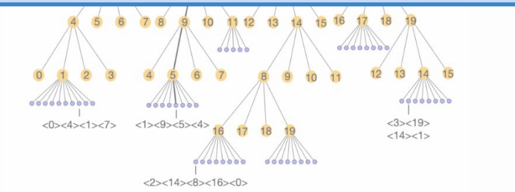
Jiang et al. Large Language Models Are Universal Recommendation Learners. arXiv: 2502.03041.

Aligning LLMs for recommendation





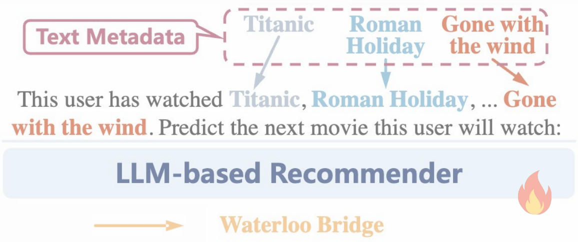
Is textual information enough for alignment?



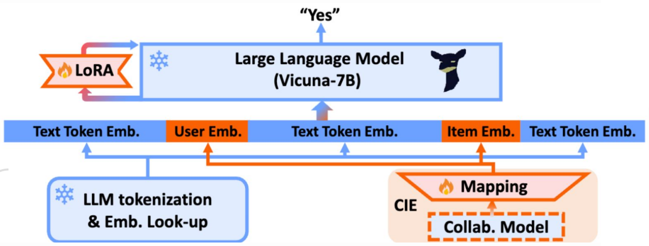
+ External item tokens

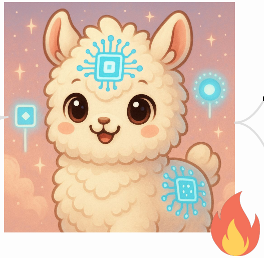
+ Multimodal information

Aligning LLMs for recommendation

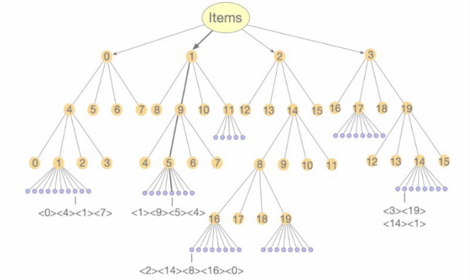


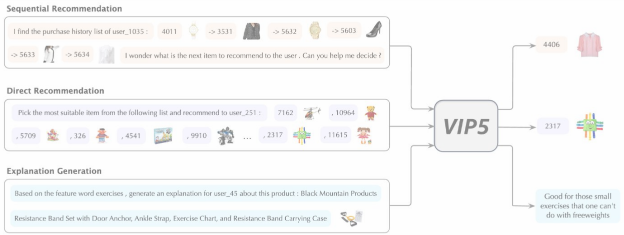
Pure text-based





+ Collaborative embeddings





+ External item tokens

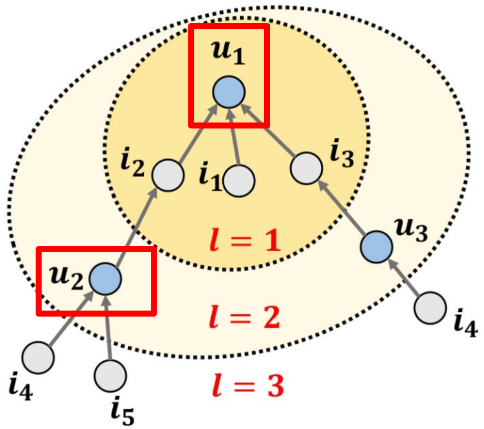
+ Multimodal information

Aligning LLMs for recommendation

(2) + Collaborative embeddings

Motivation:

Language modeling may not capture collaborative information



High-order Connectivity for u_1

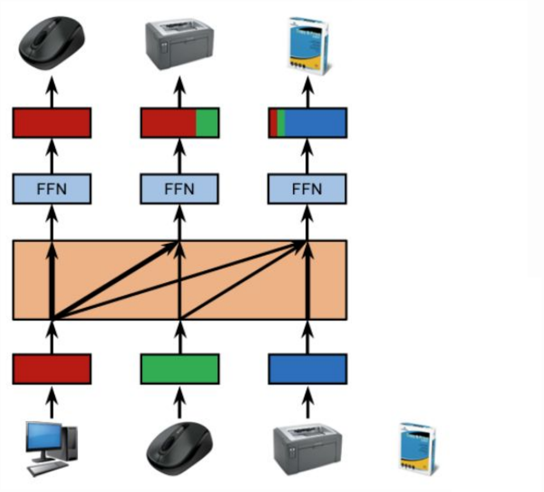
Wang et al. Neural Graph Collaborative Filtering. SIGIR 2019.

Aligning LLMs for recommendation

(2) + Collaborative embeddings

Solution:

Aligning LLMs with embeddings from traditional recommenders

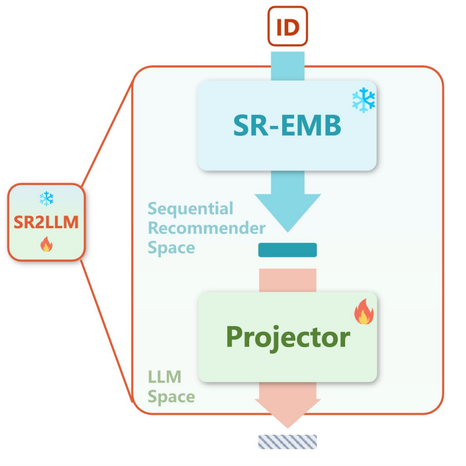


Aligning LLMs for recommendation

(2) + Collaborative embeddings (LLaRA)

- Pretrained item embeddings

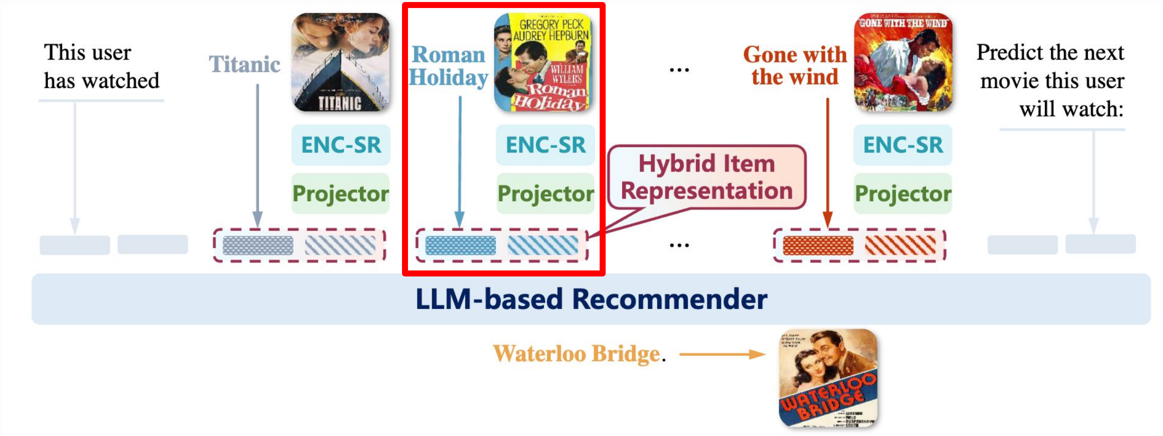
(a) Text-only prompting method.	(b) Hybrid prompting method.
Input: This user has watched Titanic [PH], Roman Holiday [PH], ... Gone with the wind [PH] in the previous. Please predict the next movie this user will watch. The movie title candidates are The Wizard of Oz [PH], Braveheart [PH],..., Waterloo Bridge [PH],... Batman & Robin [PH]. Choose only one movie from the candidates. The answer is	**Input:** This user has watched Titanic $[\mathrm{emb}_s^{14}]$, Roman Holiday $[\mathrm{emb}_s^{20}]$, ... Gone with the wind $[\mathrm{emb}_s^{37}]$ in the previous. Please predict the next movie this user will watch. The movie title candidates are The Wizard of Oz $[\mathrm{emb}_s^{5}]$, Braveheart $[\mathrm{emb}_s^{42}]$,..., Waterloo Bridge $[\mathrm{emb}_s^{20}]$,... Batman & Robin $[\mathrm{emb}_s^{19}]$. Choose only one movie from the candidates. The answer is
Output: Waterloo Bridge.	**Output:** Waterloo Bridge.



Aligning LLMs for recommendation

(2) + Collaborative embeddings (LLaRA)

+ Pretrained item embeddings

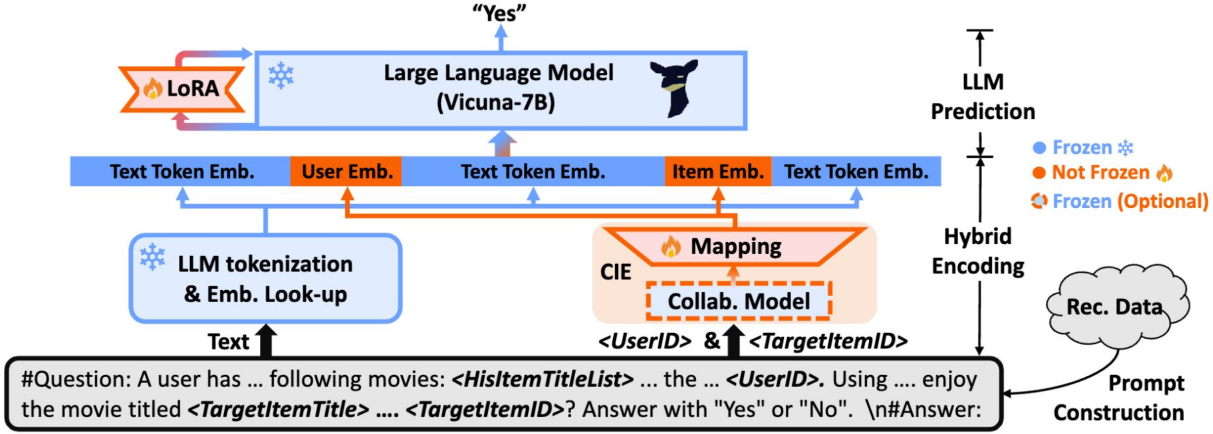


Liao et al. LLaRA: Large Language-Recommendation Assistant. SIGIR 2024.

Aligning LLMs for recommendation

(2) + Collaborative embeddings (CoLLM)

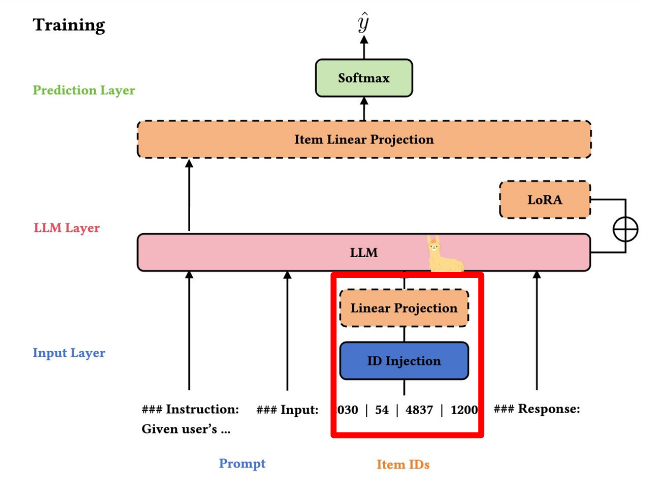
+ Pretrained item embeddings + user embeddings



Zhang et al. CoLLM: Integrating Collaborative Embeddings Into Large Language Models for Recommendation. TKDE 2025.

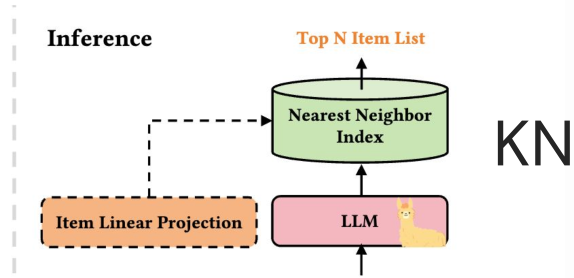
Aligning LLMs for recommendation

(2) + Collaborative embeddings (E4SRec)



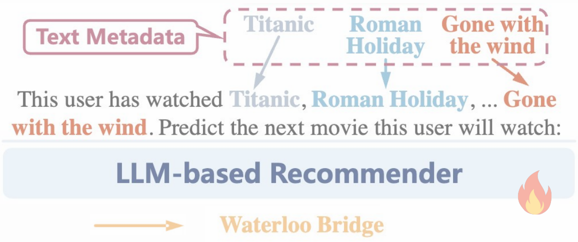
Discard text;

Collaborative embeddings only

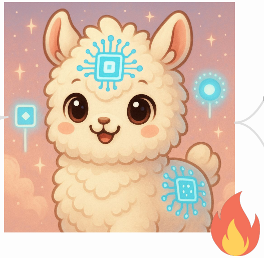


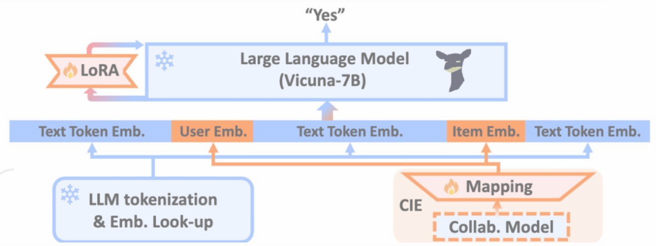
KNN for inference

Aligning LLMs for recommendation

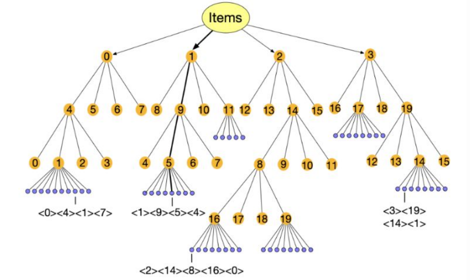


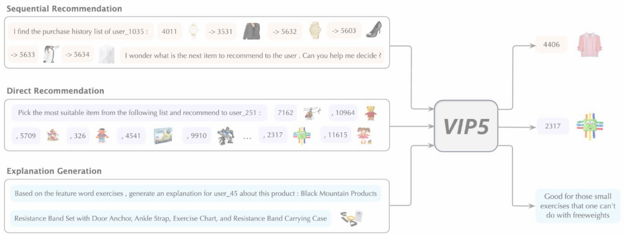
Pure text-based





+ Collaborative embeddings





+ External item tokens

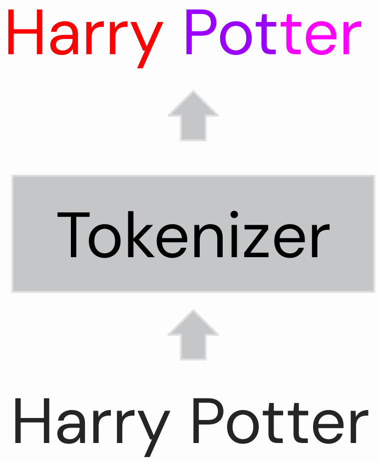
+ Multimodal information

Aligning LLMs for recommendation

(3) + External item tokens

Motivation:

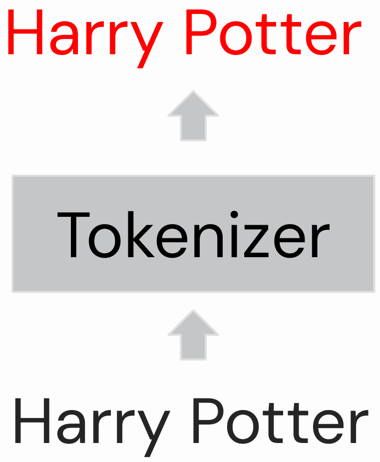
Tokens for language modeling are not optimal for recommendation.



Aligning LLMs for recommendation

(3) + External item tokens

Maybe better?

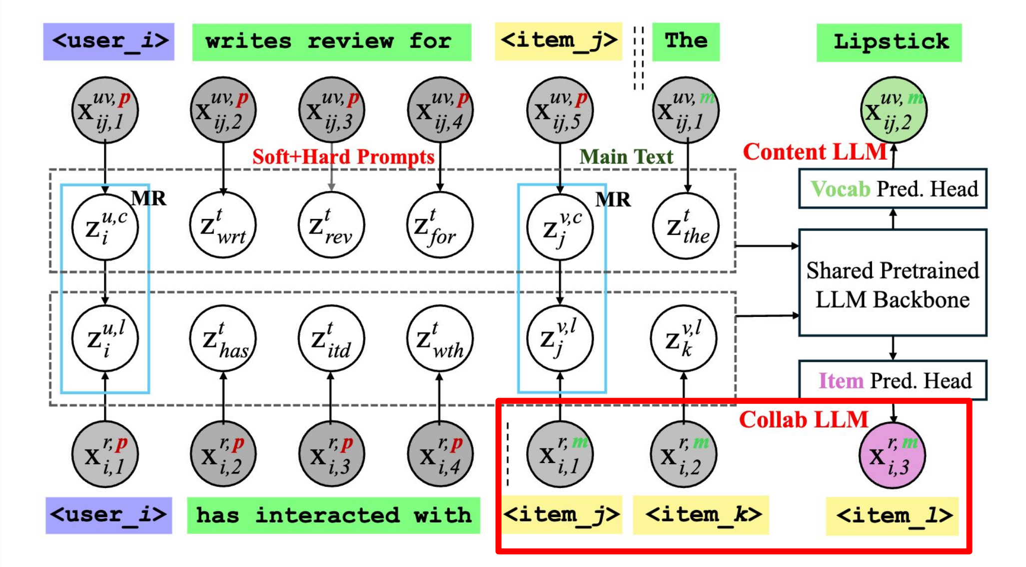


Motivation:

Tokens for language modeling are not optimal for recommendation.

Aligning LLMs for recommendation

(3) + External item tokens (CLLM4Rec)

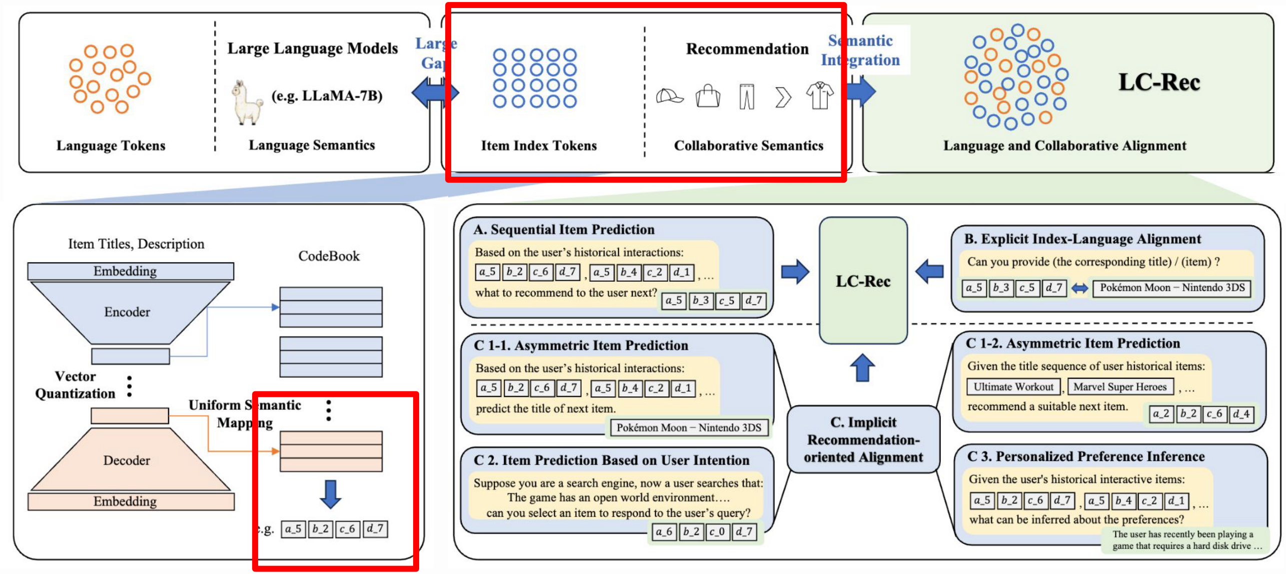


Naive approach:

One ID for each item

Aligning LLMs for recommendation

(3) + External item tokens (LC-Rec)

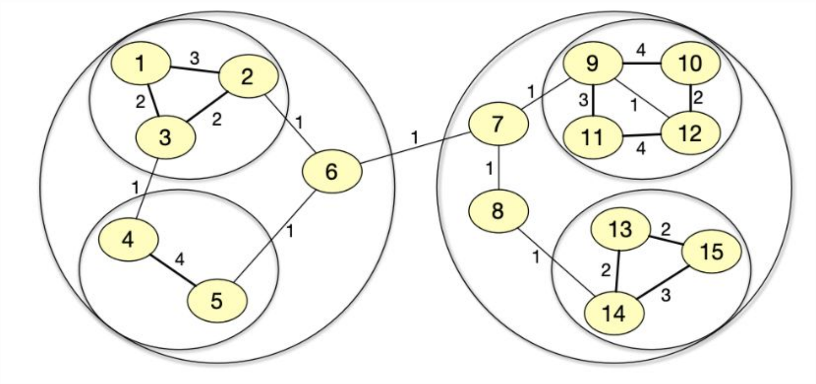


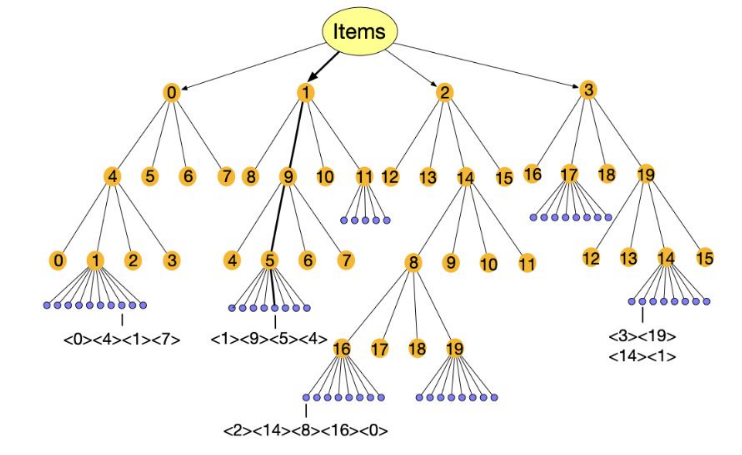
\+ Semantic IDs

(Similar items have similar IDs)

Aligning LLMs for recommendation

(3) + External item tokens

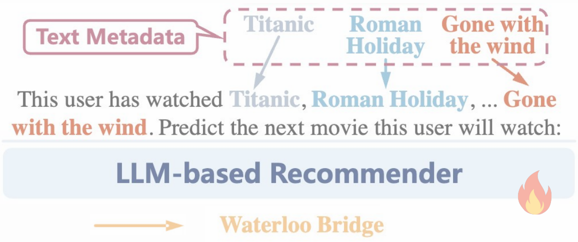




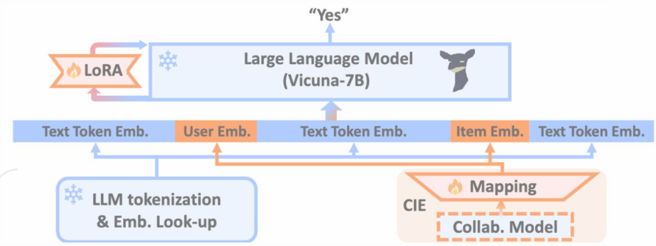
More complicated item tokens design

Hua et al. How to Index Item IDs for Recommendation Foundation Models. SIGIR-AP 2023

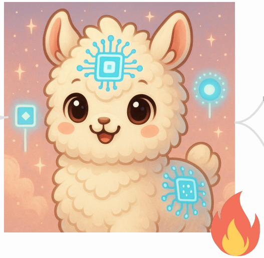
Aligning LLMs for recommendation

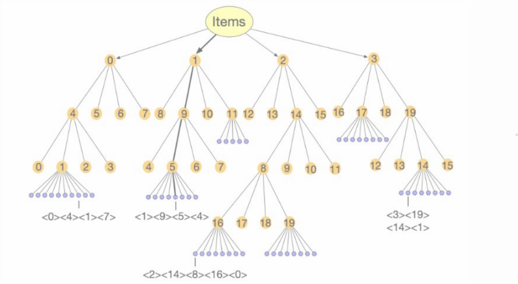


Pure text-based

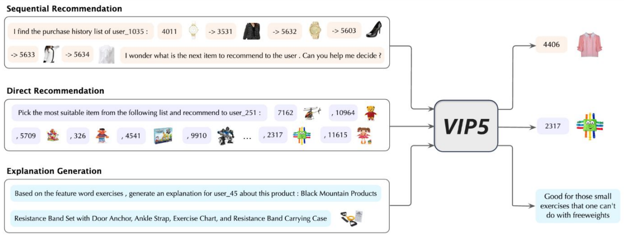


+ Collaborative embeddings





+ External item tokens



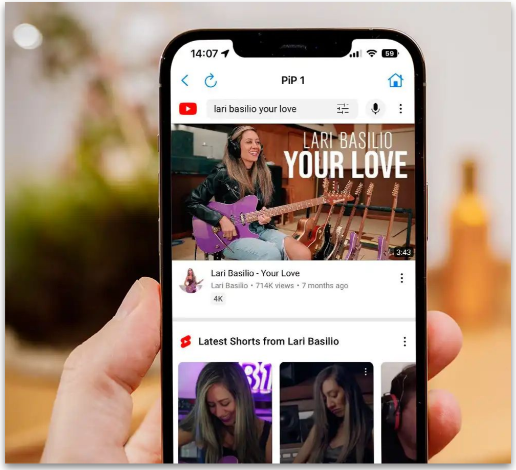
+ Multimodal information

Aligning LLMs for recommendation

(4) + Multimodal information

Motivation:

Human make decisions with multimodal information.

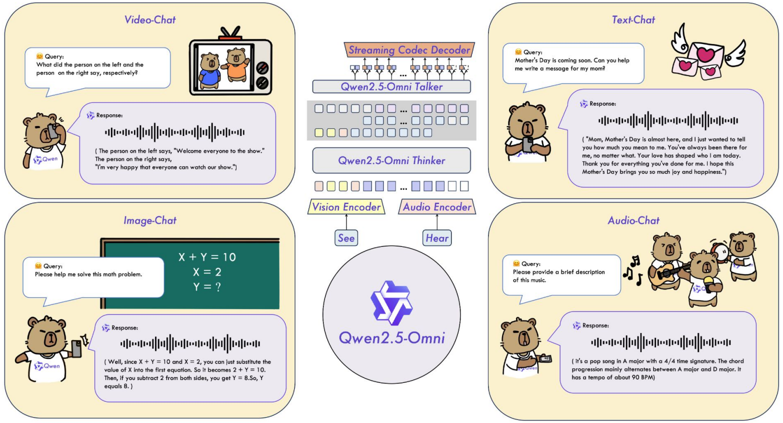


Aligning LLMs for recommendation

(4) + Multimodal information

Motivation:

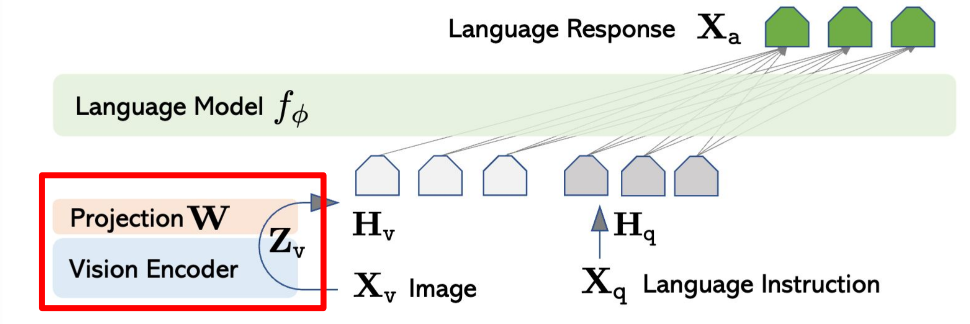
Post-trained LLM can understand multimodal information



Qwen2.5-Omni Technical Report.

Aligning LLMs for recommendation

(4) + Multimodal information

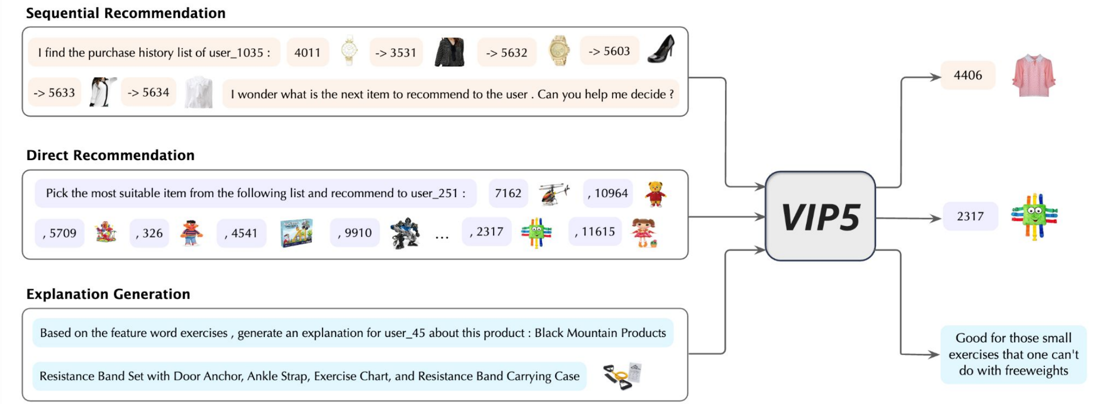


Aligning vision and language with a projector

Liu et al. Visual Instruction Tuning. NeurIPS 2023.

Aligning LLMs for recommendation

(4) + Multimodal information (VIP5)



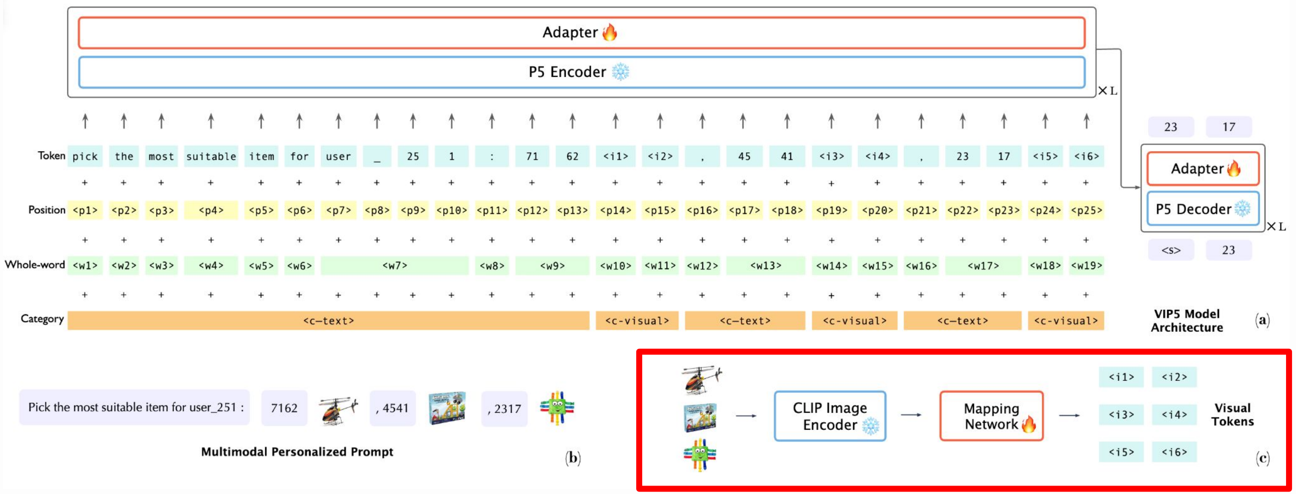
Diff between P5:

Pair text with its image

Gen et al. VIP5: Towards Multimodal Foundation Models for Recommendation. EMNLP 2023 (Findings).

Aligning LLMs for recommendation

(4) + Multimodal information (VIP5)



Alignment with projector

Gen et al. VIP5: Towards Multimodal Foundation Models for Recommendation. EMNLP 2023 (Findings).

Aligning LLMs for recommendation

(4) + Multimodal information (VIP5)

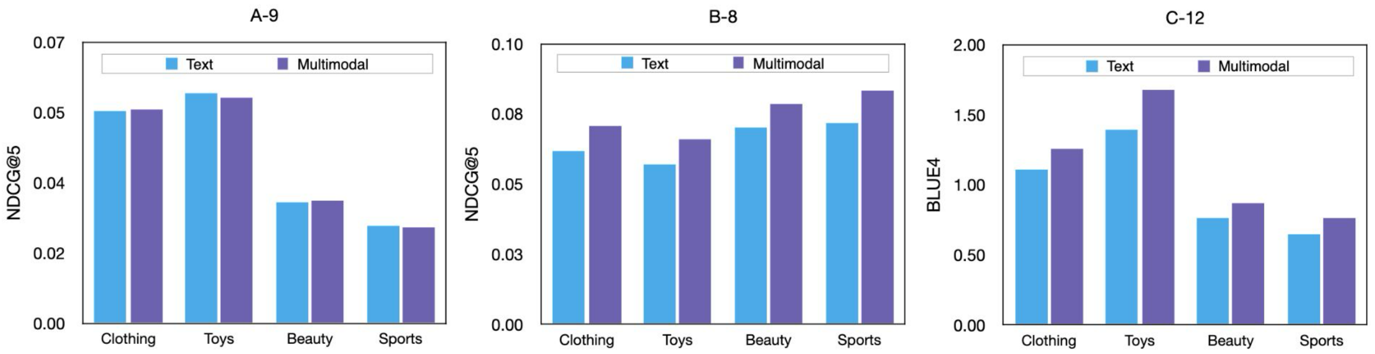
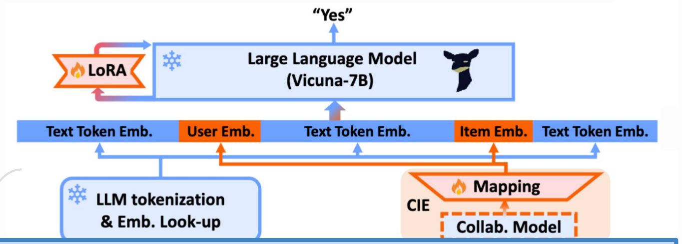


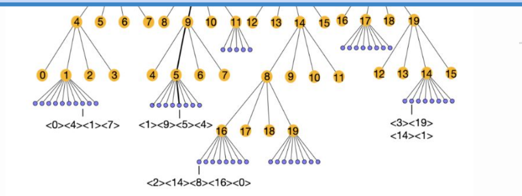
Figure 3: Performance comparison between text-based prompt and multimodal prompt.

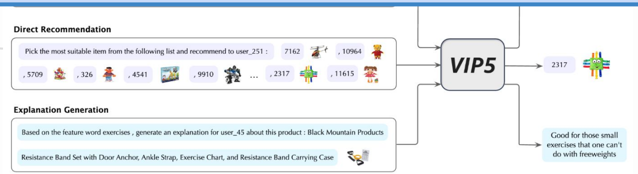
Multimodal information is important

Aligning LLMs for recommendation



Information tailored for recommendation matters!





+ External item tokens

+ Multimodal information

Part 1: LLM as Sequential Recommender

(i) Early efforts: Pretrained LLMs for recommendation;
(ii) Aligning LLMs for recommendation;
(iii) **Training** objective & **inference**

Training objective

(1) Supervised finetuning (SFT)

I have watched Titanic, Roman Holiday, … Gone with the wind. Predict the next movie I will watch:

Training objective

(1) Supervised finetuning (SFT)

I have watched Titanic, Roman Holiday, … Gone with the wind. Predict the next movie I will watch:

Waterloo Bridge.

Prediction

Training objective

(1) Supervised finetuning (SFT)

I have watched Titanic, Roman Holiday, … Gone with the wind. Predict the next movie I will watch:

Waterloo Bridge.

Prediction

Training objective

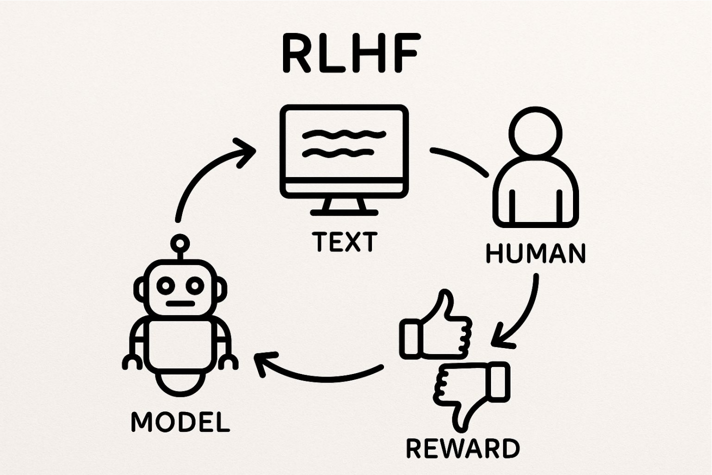
(1) Supervised finetuning (SFT)

$$\mathcal{L}_{\mathrm{SFT}}(\theta) = -\mathbb{E}_{(x,y)\sim\mathcal{D}}\left[\sum_{t=1}^{T} \log \boxed{P_\theta(y_t \mid y_{<t})}\right]$$

Always predict the next token

Training objective

(2) Preference learning



LLMs are trained to align human preferences

Recommendation is about user preferences

Training objective

(2) Preference learning

I have watched Titanic, Roman Holiday, … Gone with the wind. Predict the next movie I will watch:

Waterloo Bridge > Harry Potter

Training objective

(2) Preference learning

$$\mathcal{L}_{\mathrm{DPO}} = -\mathbb{E}_{(x_u, e_p, e_d)}\left[\log \sigma\left(\beta \log \frac{\pi_\theta(e_p|x_u)}{\pi_{\mathrm{ref}}(e_p|x_u)} - \beta \log \frac{\pi_\theta(e_d|x_u)}{\pi_{\mathrm{ref}}(e_d|x_u)}\right)\right],$$

Direct Preference Optimization!

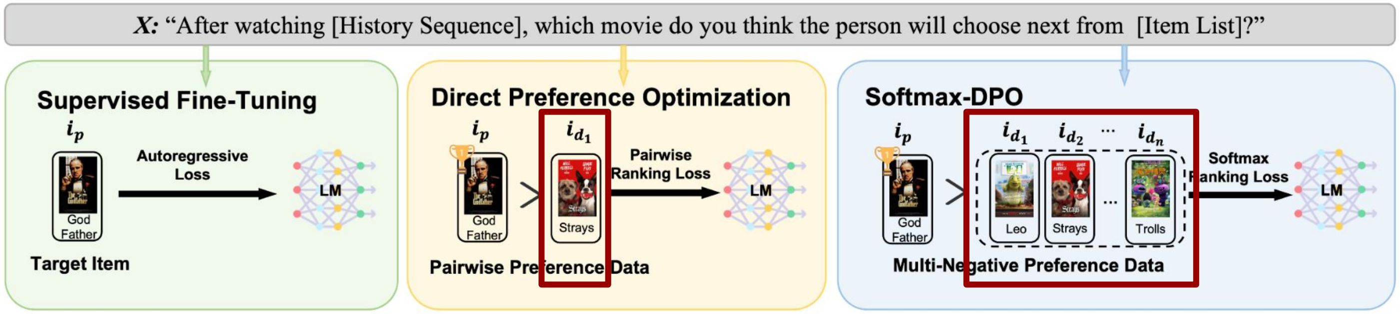
I have watched Titanic, Roman Holiday, … Gone with the wind. Predict the next movie I will watch:

Waterloo Bridge > Harry Potter

Rafailov et al. Direct Preference Optimization: Your Language Model is Secretly a Reward Model. NeurIPS 2023.

Training objective

(2) Preference learning

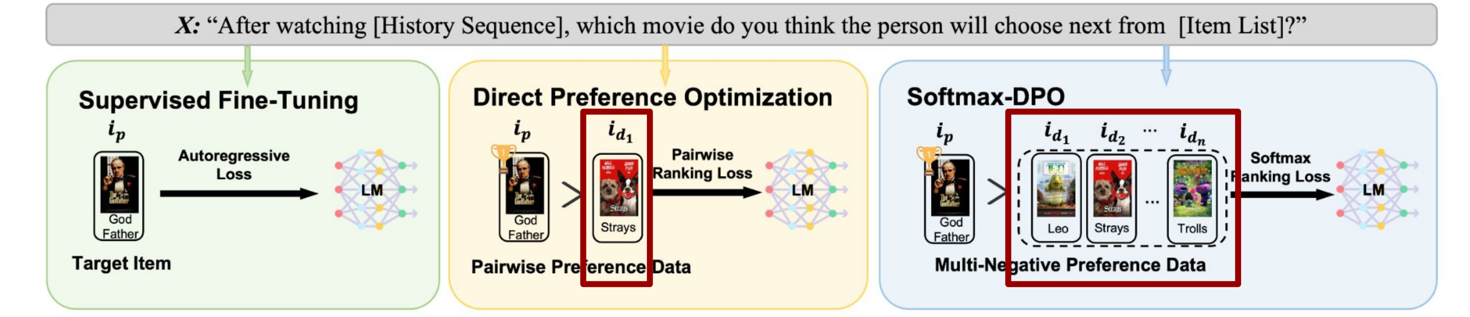


Single negative → Multiple negatives

Chen et al. On Softmax Direct Preference Optimization for Recommendation. NeurIPS 2024.

Training objective

(2) Preference learning



$$\mathcal{L}_{\mathrm{S-DPO}}(\pi_\theta;\pi_{\mathrm{ref}}) = -\mathbb{E}_{(x_u,e_p,\mathcal{E}_d)\sim\mathcal{D}}\left[\log\sigma\left(-\log\sum_{e_d\in\mathcal{E}_d}\exp\left(\beta\log\frac{\pi_\theta(e_d|x_u)}{\pi_{\mathrm{ref}}(e_d|x_u)}-\beta\log\frac{\pi_\theta(e_p|x_u)}{\pi_{\mathrm{ref}}(e_p|x_u)}\right)\right)\right].$$

Chen et al. On Softmax Direct Preference Optimization for Recommendation. NeurIPS 2024.

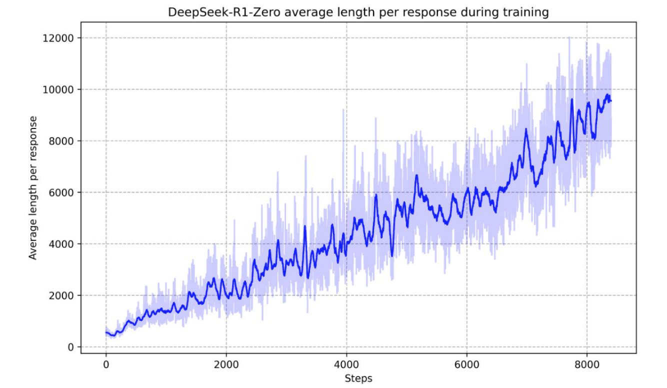
Training objective

(3) Reinforce learning

Emergent reasoning capabilities through RL

$$\mathcal{J}_{GRPO}(\theta) = \mathbb{E}[q \sim P(Q), \{o_i\}_{i=1}^{G} \sim \pi_{\theta_{old}}(O|q)]$$
$$\frac{1}{G}\sum_{i=1}^{G}\left(\min\left(\frac{\pi_\theta(o_i|q)}{\pi_{\theta_{old}}(o_i|q)}A_i, \text{clip}\left(\frac{\pi_\theta(o_i|q)}{\pi_{\theta_{old}}(o_i|q)}, 1-\varepsilon, 1+\varepsilon\right)A_i\right) - \beta\mathbb{D}_{KL}\left(\pi_\theta||\pi_{ref}\right)\right), \quad (1)$$

$$\mathbb{D}_{KL}\left(\pi_\theta||\pi_{ref}\right) = \frac{\pi_{ref}(o_i|q)}{\pi_\theta(o_i|q)} - \log\frac{\pi_{ref}(o_i|q)}{\pi_\theta(o_i|q)} - 1, \quad (2)$$



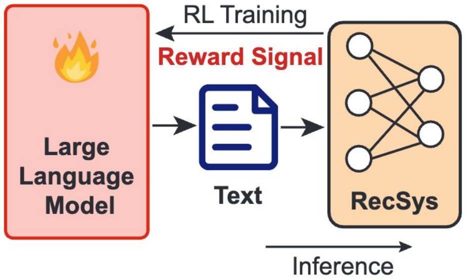
Question: If $a > 1$, then the sum of the real solutions of $\sqrt{a - \sqrt{a + x}} = x$ is equal to

Response: <think>
To solve the equation $\sqrt{a - \sqrt{a + x}} = x$, let's start by squaring both $\cdots$
$\left(\sqrt{a - \sqrt{a + x}}\right)^2 = x^2 \implies a - \sqrt{a + x} = x^2.$
Rearrange to isolate the inner square root term:
$(a - x^2)^2 = a + x \implies a^2 - 2ax^2 + (x^2)^2 = a + x \implies x^4 - 2ax^2 - x + (a^2 - a) = 0$
$\ldots$
Wait, wait. Wait. That's an aha moment I can flag here.
Let's reevaluate this step-by-step to identify if the correct sum can be $\cdots$
We started with the equation:
$\sqrt{a - \sqrt{a + x}} = x$
First, let's square both sides:
$a - \sqrt{a + x} = x^2 \implies \sqrt{a + x} = a - x^2$
Next, I could square both sides again, treating the equation: $\cdots$
$\ldots$

DeepSeek-AI et al. DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning. arXiv: 2501.12948.

Training objective

(3) Reinforce learning



$$\max_{\theta} \mathbb{E}_{s\sim p(s),a\sim\pi_\theta(a|s)} \boxed{[f(a|s)]}.$$

Maximize the reward from recommender system

Prompt Template for Rec-R1 + Dense Retriever (Product Search)

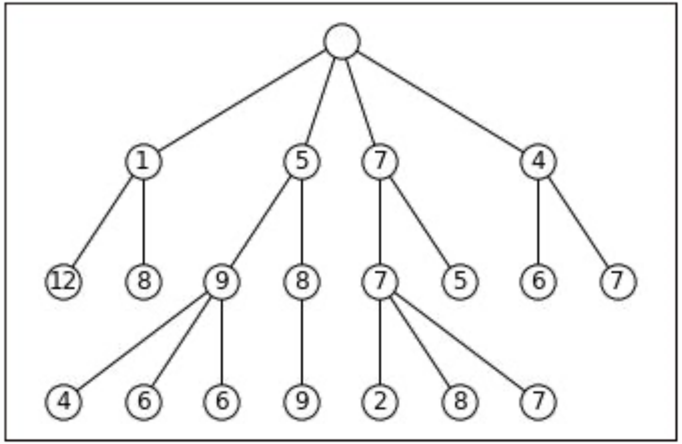
```
You are an expert in generating queries for dense retrieval. Given a customer
query, your task is to retain the original query while expanding it with
additional semantically relevant information, retrieve the most relevant
products, ensuring they best meet customer needs. If no useful expansion is
needed, return the original query as is.

Below is the query:
``` {user_query} ```

<|im_start|>system
You are a helpful AI assistant. You first think about the reasoning process
in the mind and then provide the user with the answer.
<|im_end|>
<|im_start|>user
[PROMPT as above]
Show your work in <think>\think> tags. Your final response must be in JSON
format within <answer>\answer> tags. For example,
<answer>
{ "query": xxx }
</answer>.
<|im_end|>
<|im_start|>assistant
Let me solve this step by step.
<think>
```
```

Inference

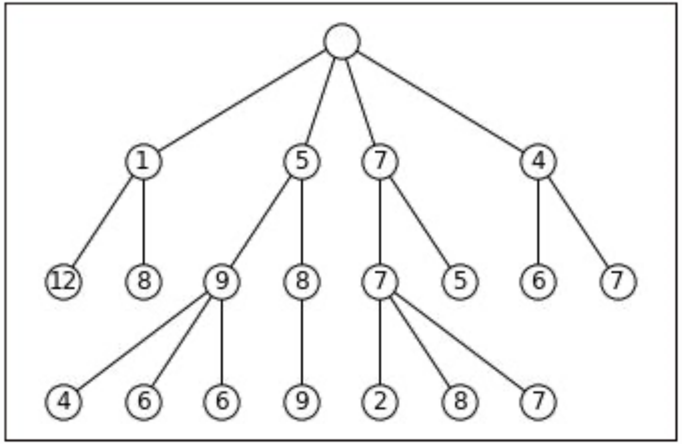
(1) Beam Search



Generating answers with the top-k highest scored beams

Inference

(1) Beam Search



It may generate invalid items

In RecSys :
No Hallucination permitted!

Inference

(2) Constrained Beam Search

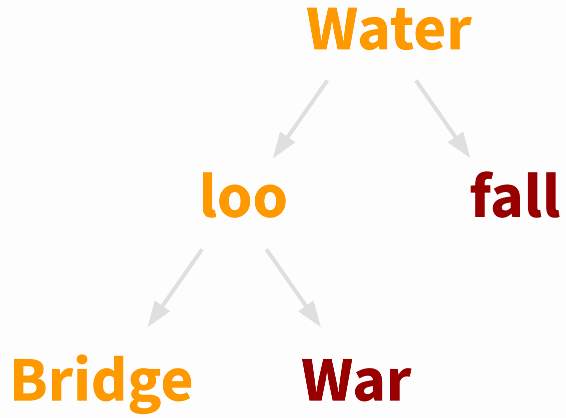
Valid items:
Waterloo Bridge, Waterfall Story, and Waterloo War

How to make the generated items always valid?

Inference

(2) Constrained Beam Search

Valid items:
Waterloo Bridge, Waterfall Story, and Waterloo War



Constrained search tree

Inference

(2) Constrained Beam Search

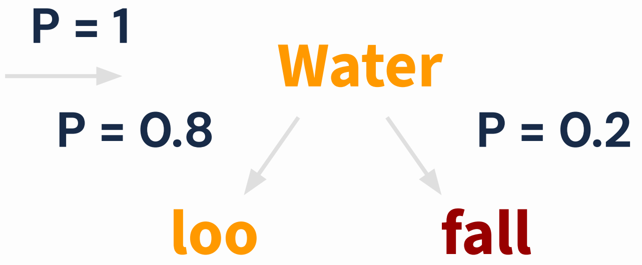
I have watched Titanic, Roman Holiday, … Gone with the wind. Predict the next movie I will watch:

P = 1 → **Water**

Inference

(2) Constrained Beam Search

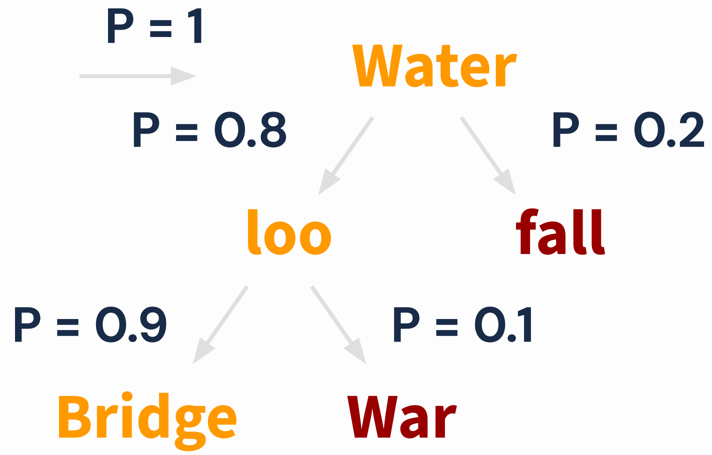
I have watched Titanic, Roman Holiday, … Gone with the wind. Predict the next movie I will watch:



Inference

(2) Constrained Beam Search

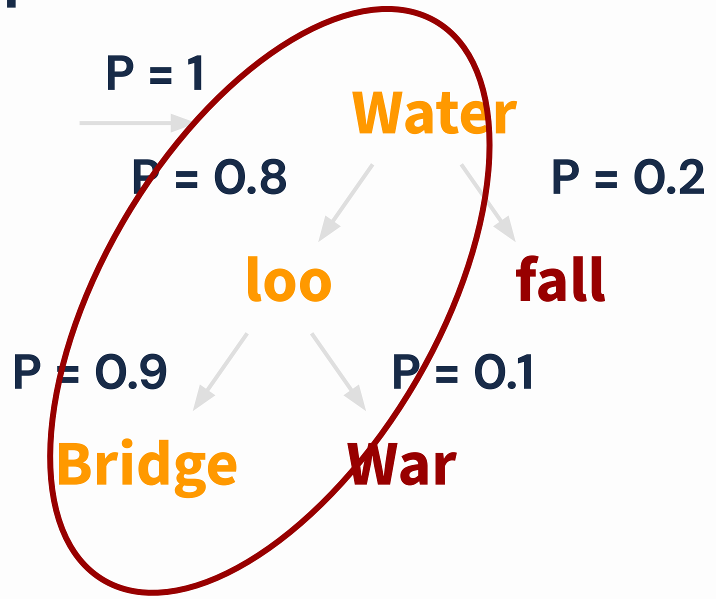
I have watched Titanic, Roman Holiday, … Gone with the wind. Predict the next movie I will watch:



Inference

(2) Constrained Beam Search

I have watched Titanic, Roman Holiday, …
Gone with the wind. Predict the next movie
I will watch:



Valid Item!

Inference

(3) Special design

$$\mathcal{S}(h_{\leq t}) = \mathcal{S}(h_{\leq t-1}) + \log(p(h_t|x, h_{\leq t-1})),$$

$$\mathcal{S}(h) = \mathcal{S}(h)\boxed{/h_L^{\alpha},}$$

Length penalty in beam search;
Human does not like over long sentences.

Redundant for recommendation

Bao et al. Decoding Matters: Addressing Amplification Bias and Homogeneity Issue for LLM-based Recommendation. EMNLP 2024.

Inference

(3) Special design

$$\mathcal{S}(h_{\leq t}) = \mathcal{S}(h_{\leq t-1}) + \log(p(h_t | x, h_{\leq t-1})),$$

$$\mathcal{S}(h) = \mathcal{S}(h) / \cancel{L},$$

Remove length penalty

	Instruments	Books	CDs	Sports	Toys	Games
Baseline	0.1062	0.0308	0.0956	0.1171	0.0965	0.0610
D^3	**0.1111**	**0.0354**	**0.1190**	**0.1215**	**0.1025**	**0.0767**
- RLN	0.1093	0.0353	0.1000	0.1200	0.0975	0.0659
- TFA	0.1086	0.0309	0.1115	0.1192	0.1006	0.0732

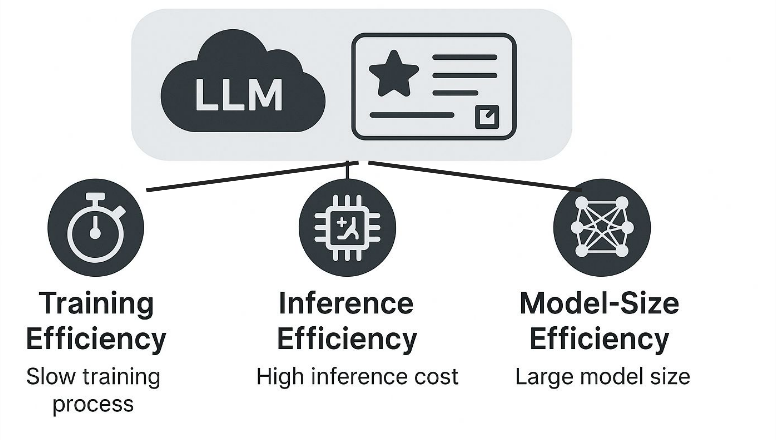
Imp when removing

Part 1: LLM as Sequential Recommender

(i) Early efforts: Pretrained LLMs for recommendation;
(ii) Aligning LLMs for recommendation;
(iii) Training objective & inference
(iiii) **Efficiency**

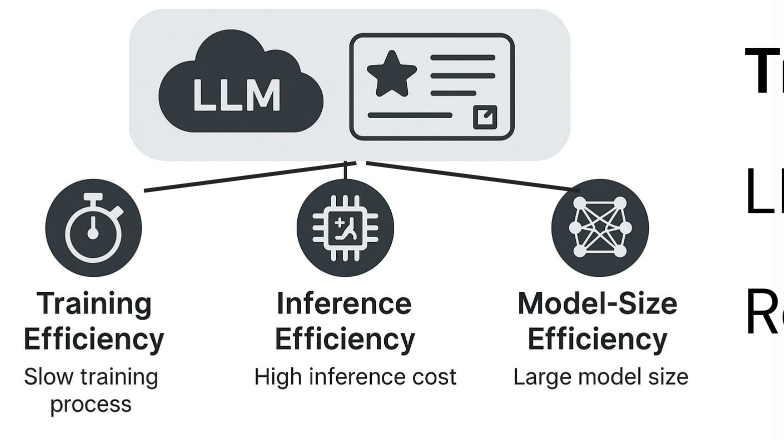
Efficiency

A crucial question in real-world deployment



Efficiency

A crucial question in real-world deployment



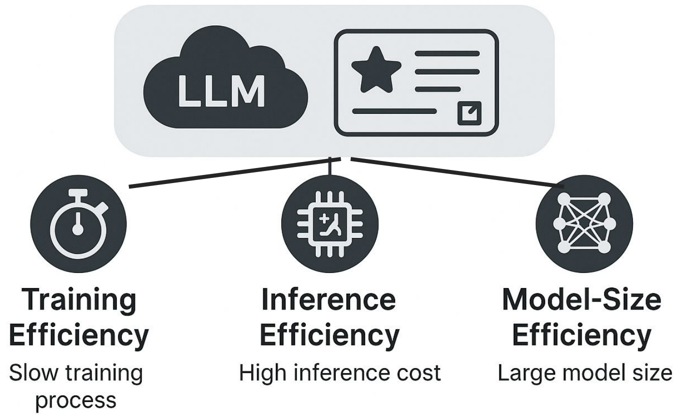
Training efficiency:

LLM: update by months

Recommender: update by hours

Efficiency

A crucial question in real-world deployment



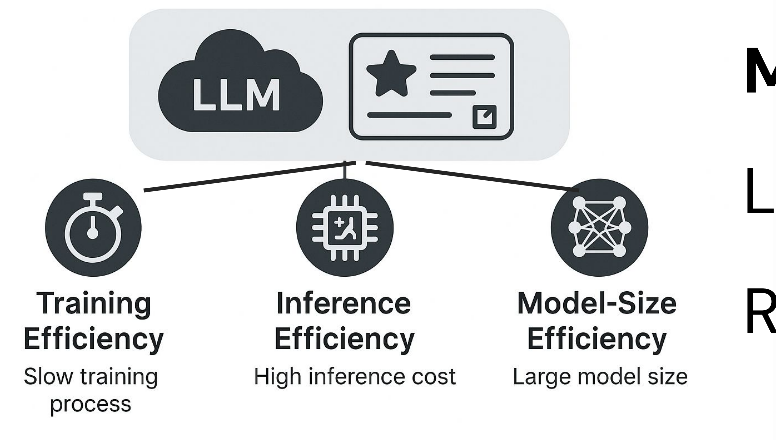
Inference efficiency:

LLM: wait for seconds

Recommender: wait for milliseconds

Efficiency

A crucial question in real-world deployment



Model-size efficiency:

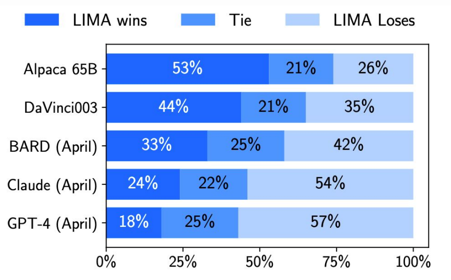
LLM: serve for millions

Recommender: serve for billions

Efficiency

(1) Training efficiency

Source	#Examples
Training	
Stack Exchange (STEM)	200
Stack Exchange (Other)	200
wikiHow	200
Pushshift r/WritingPrompts	150
Natural Instructions	50
Paper Authors (Group A)	200
Dev	
Paper Authors (Group A)	50
Test	
Pushshift r/AskReddit	70
Paper Authors (Group B)	230



Less is more for alignment

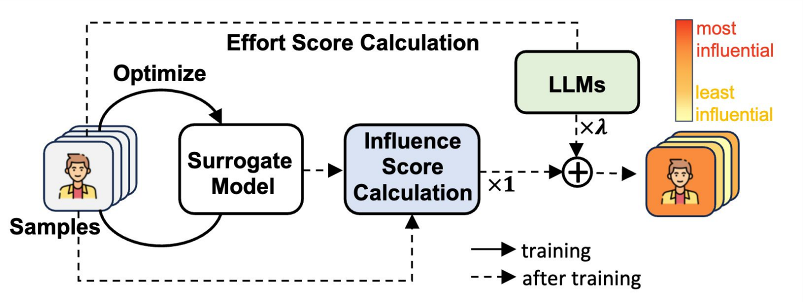
1k high quality examples ->

Surpass large scale training

Zhou et al. LIMA: Less Is More for Alignment. NeurIPS 2023.

Efficiency

(1) Training efficiency



	Games				
	R@10↑	R@20↑	N@10↑	N@20↑	Time↓
Full	0.0169	0.0233	0.0102	0.0120	36.87h
DEALRec	0.0181	0.0276	0.0115	0.0142	1.67h
% Improve.	**7.10%**	**18.45%**	**12.75%**	**18.33%**	**-95.47%**

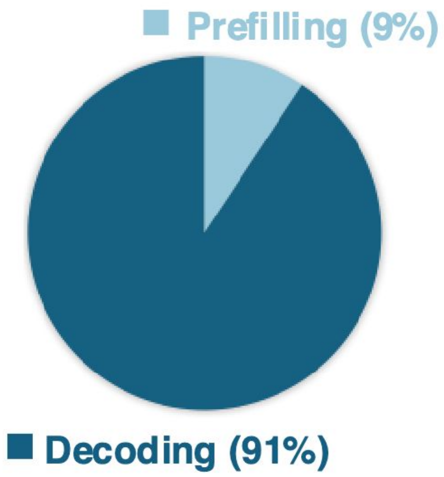
Select the most informative examples ->

Reducing 95% training time

Lin et al. Data-efficient Fine-tuning for LLM-based Recommendation. SIGIR 2024.

Efficiency

(2) Inference efficiency



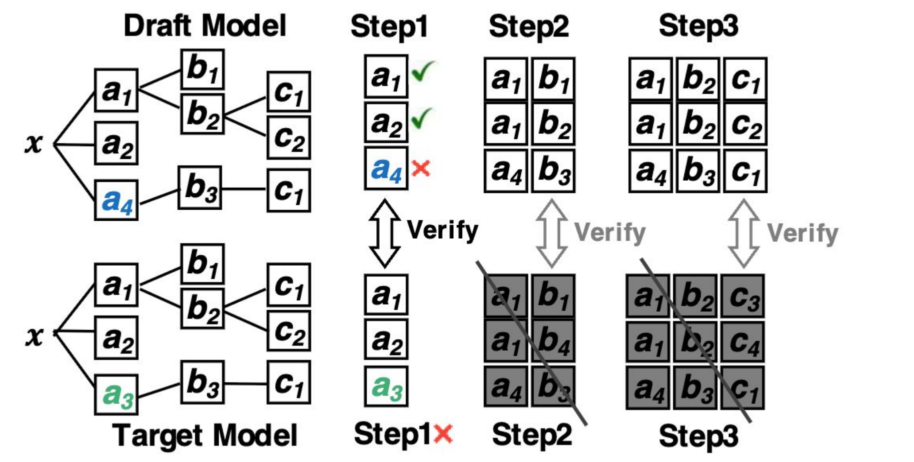
Autoregressive paradigm in LLM

-> huge time on the decoding stage

Lin et al. Efficient Inference for Large Language Model-based Generative Recommendation. ICLR 2025.

Efficiency

(2) Inference efficiency



N-to-K Verification of SD with Beam Search (N=K=3)

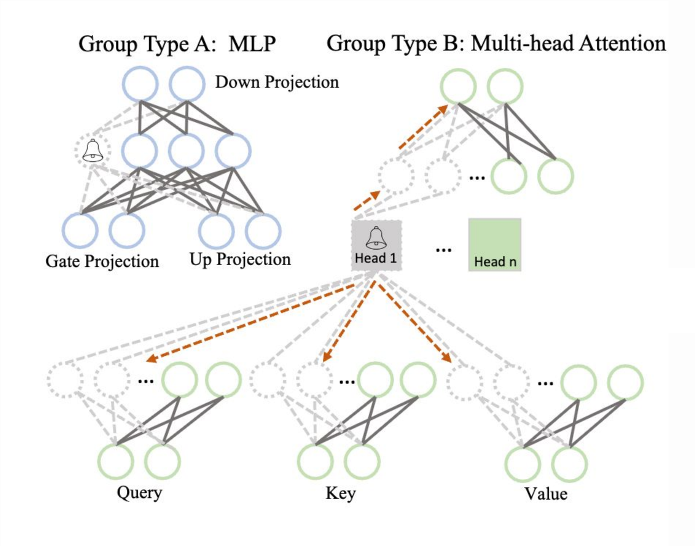
Speculative decoding:

Decoder acceleration with a small-size draft model

Lin et al. Efficient Inference for Large Language Model-based Generative Recommendation. ICLR 2025.

Efficiency

(3) Model-size efficiency – Pruning



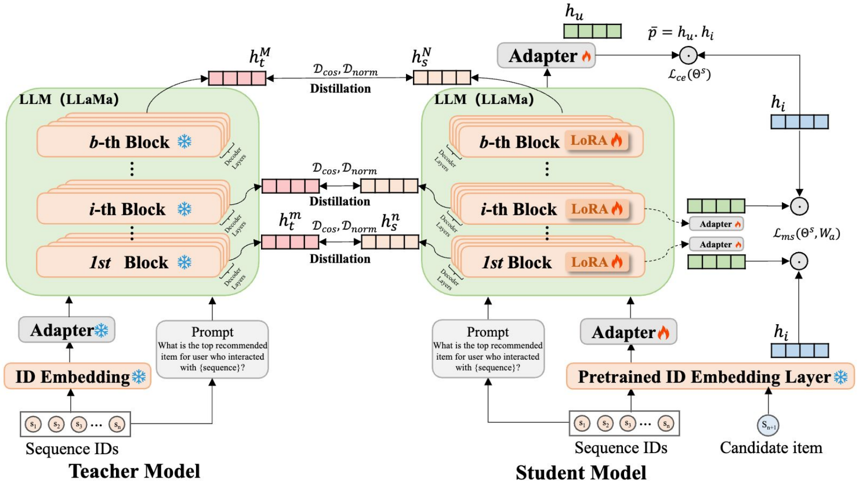
Similar performance with 0.6% parameters

Method	#Params	Tasks		
		TNEWS↑	IFLYTEK↑	CSL↑
M6-base	327M	0.598	0.631	0.852
ALBERT-zh-base	12M	0.550	0.564	0.785
M6-Edge	10M	**0.552**	**0.586**	**0.831**
ALBERT-zh-tiny	4M	0.534	0.488	0.750
M6-Edge, Pruned	2M	**0.537**	**0.559**	**0.798**

Ma et al. LLM-Pruner: On the Structural Pruning of Large Language Models. NeurIPS 2023
Cui et al. M6-Rec: Generative Pretrained Language Models are Open-Ended Recommender Systems. arXiv: 2205.08084.

Efficiency

(3) Model-size efficiency – Distillation



SLM learns from LLM

With Hard label + soft label

Efficiency

(3) Model-size efficiency – Distillation

Table 3: Experimental results (%) on the Music and Sport dataset.

Model	Music				Sport				Rank
	HR@1	HR@5	NDCG@5	MRR	HR@1	HR@5	NDCG@5	MRR	
Caser	0.71	3.28	1.96	2.29	1.05	3.75	2.39	2.84	13.50
GRU4Rec	1.89	3.22	2.57	3.08	5.26	7.75	6.52	7.08	10.13
BERT4Rec	2.10	3.16	2.64	3.11	4.81	6.70	5.79	6.26	10.63
SASRec	1.82	5.72	3.79	4.51	4.70	8.43	6.59	7.24	8.75
HGN	2.01	5.49	3.82	4.17	3.42	6.24	4.83	5.30	10.50
LightSANs	1.05	4.06	2.54	3.00	5.18	8.94	7.07	7.72	8.25
S^3-Rec	2.48	7.37	4.94	4.68	4.14	8.49	6.89	7.35	6.88
DuoRec	1.84	4.50	3.19	3.04	4.13	8.81	7.03	6.64	9.13
MAERec	2.19	6.35	4.67	3.96	4.01	8.35	6.65	6.98	8.63
Open-P5	4.35	8.12	6.74	-	5.49	8.50	6.92	-	5.33
E4SRec	5.62	9.29	7.50	7.98	6.40	9.67	8.05	8.70	1.75
$E4SRec_8$	5.46	8.86	7.21	7.74	5.48	8.63	7.06	7.76	3.63
$E4SRec_4$	5.33	8.75	7.08	7.59	5.41	8.65	7.04	7.72	4.50
$SLMRec_{4\leftarrow 8}$	**5.72**	**9.15**	**7.48**	**8.03**	**6.62**	**9.83**	**8.25**	**8.89**	**1.25**

Method	Tr time(h)	Inf time(h)	Tr params (B)	Inf params (B)
$\text{Open-P5}_{\text{LLaMa}}$	0.92	4942	0.023	7.237
E4SRec	3.95	0.415	0.023	6.631
$\textbf{SLMRec}_{4\leftarrow 8}$	0.60	0.052	0.003	0.944

Reduced model-size;

Reduced inference time

Xu et al..SLMRec: Distilling Large Language Models into Small for Sequential Recommendation. ICLR 2025.

Part 1: LLM as Sequential Recommender

(1) Early efforts: pretrained LLMs for rec

(2) Aligning LLMs for recommendation

- Pure text-based
- External item tokens
- Collaborative embeddings
- Multimodal information

(3) Training objective & inference

Training: SFT, DPO, RL; **Inference**: (constrained) beam search

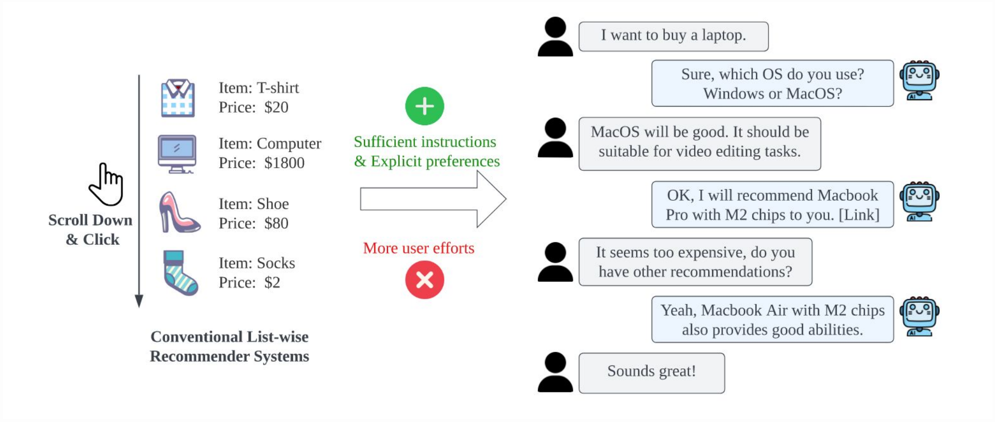
(4) Efficiency

Data efficiency; Inference efficiency; Model-size efficiency

Part 2: LLM as Conversational Recommender

Conversational Recommender System (CRS)

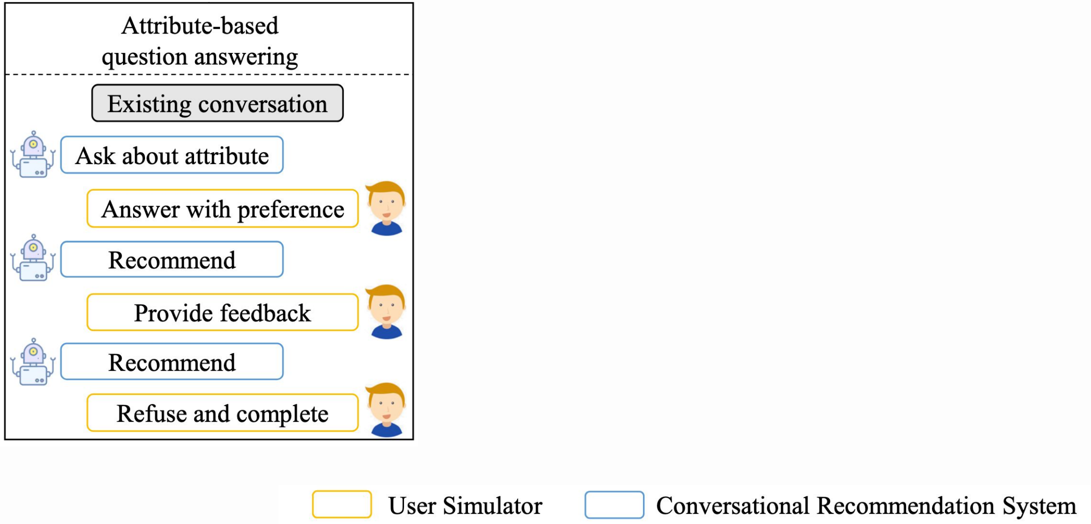
- Recommendations with multiple turns conversation
- Interactive; engaging users in the loop



Chen et al. All Roads Lead to Rome: Unveiling the Trajectory of Recommender Systems Across the LLM Era. arXiv.2407.10081

Paradigms of CRS before the era of LLM

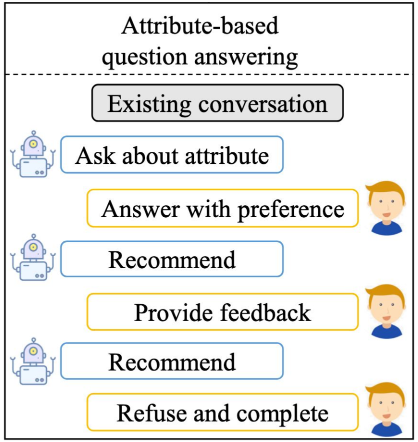
Attribute-based



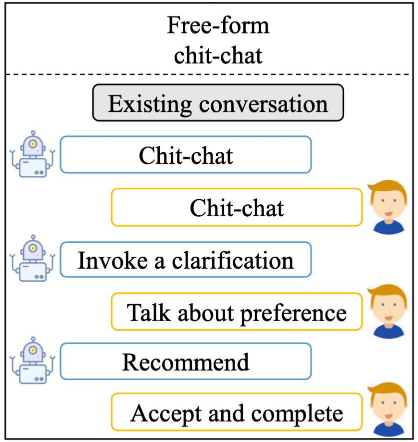
Wang et al. Rethinking the Evaluation for Conversational Recommendation in the Era of Large Language Models. EMNLP 2023.

Paradigms of CRS before the era of LLM

Attribute-based



Free-form



User Simulator   Conversational Recommendation System

Wang et al. Rethinking the Evaluation for Conversational Recommendation in the Era of Large Language Models. EMNLP 2023.

Paradigms of CRS before the era of LLM

Features: Task-specific conversational recommenders, trained on limited conversation data.

Paradigms of CRS before the era of LLM

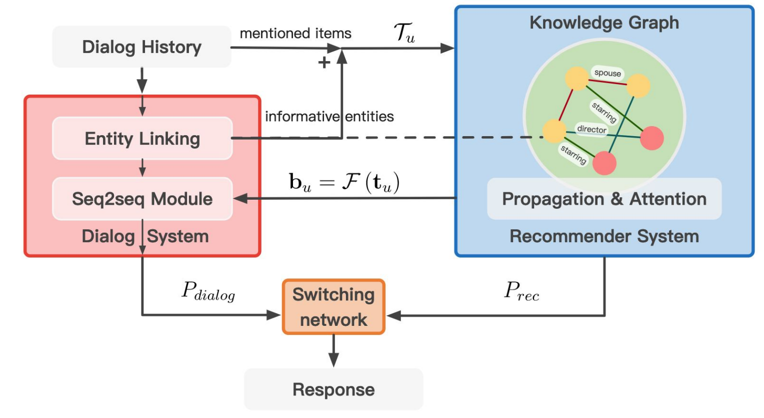
Features: Task-specific conversational recommenders, trained on limited conversation data.

- Lack of world knowledge.
- Requirement of complicated strategies.
- Incompatible natural language generation abilities.
- Lack of generalization capabilities.

Paradigms of CRS before the era of LLM

Traditional CRS: KBRD

- End-to-end conversational recommender system
- Switching between conversation and recommendation
- External knowledge from knowledge graph



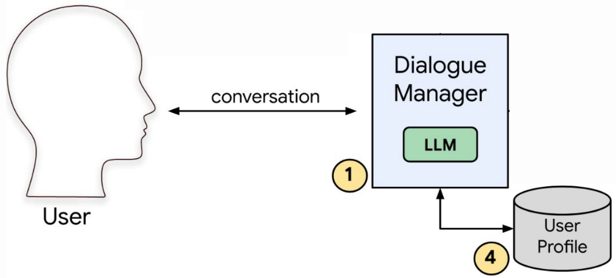
Chen et al. Towards Knowledge-Based Recommender Dialog System. EMNLP 2019.

Example

LLM as conversational recommender

LLM as Conversational Recommender

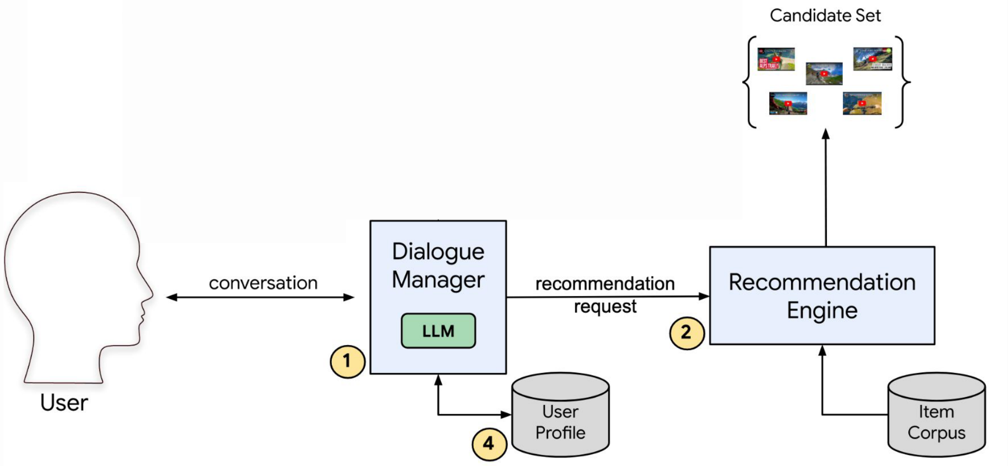
Framework (RecLLM)



Conversation with users via LLMs

LLM as Conversational Recommender

Framework (RecLLM)

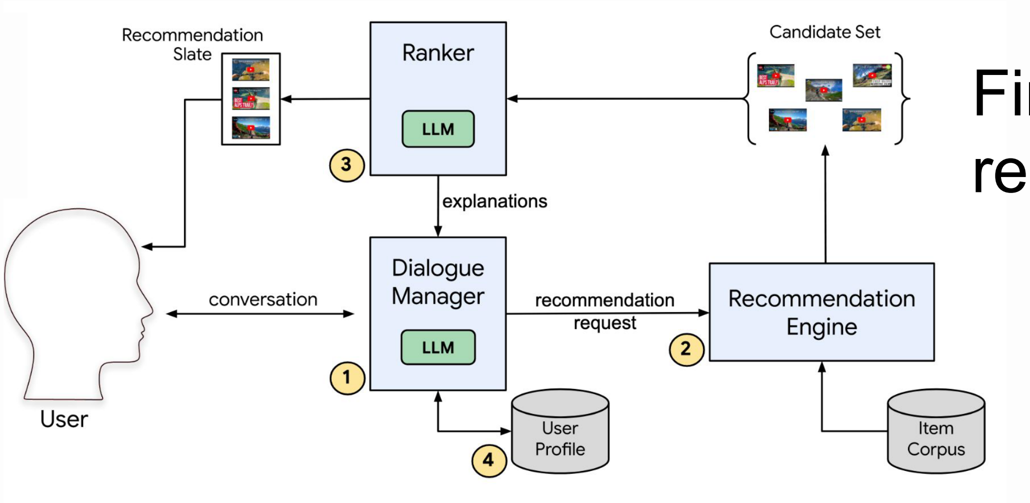


Recommendation via tools

Friedman et al. Leveraging Large Language Models in Conversational Recommender Systems. arXiv:2305.07961.

LLM as Conversational Recommender

Framework (RecLLM)



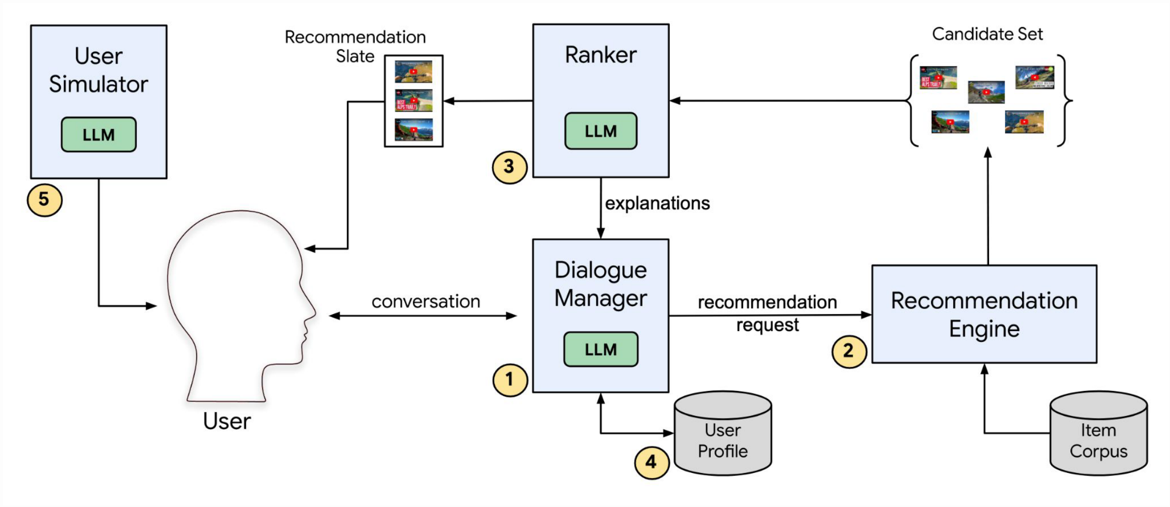
Fine-grained reranking via LLMs

Friedman et al. Leveraging Large Language Models in Conversational Recommender Systems. arXiv:2305.07961.

LLM as Conversational Recommender

Framework (RecLLM)

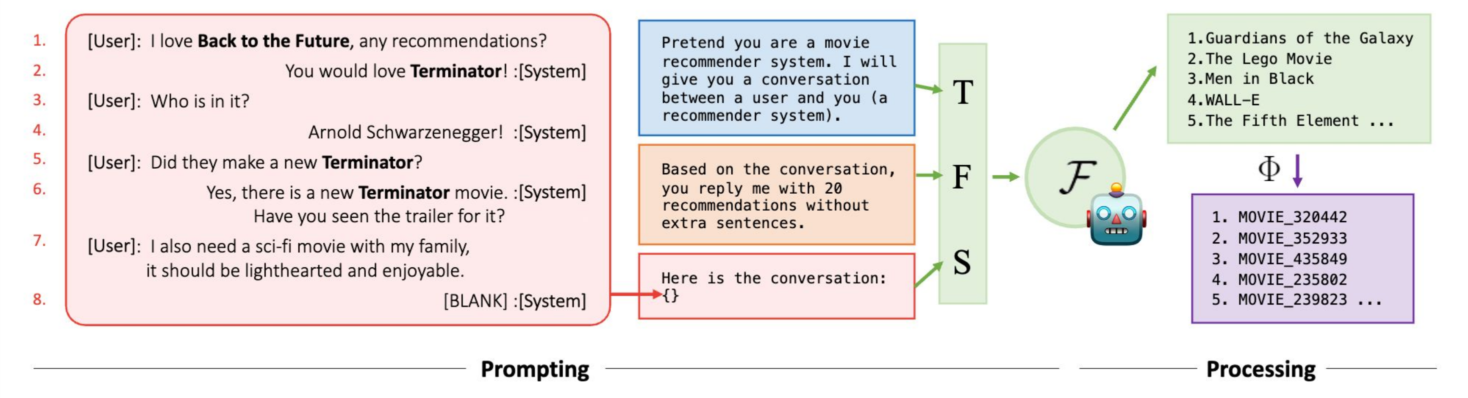
Evaluation via LLMs



Friedman et al. Leveraging Large Language Models in Conversational Recommender Systems. arXiv:2305.07961.

LLM as Conversational Recommender

LLMs as zero-shot CRS

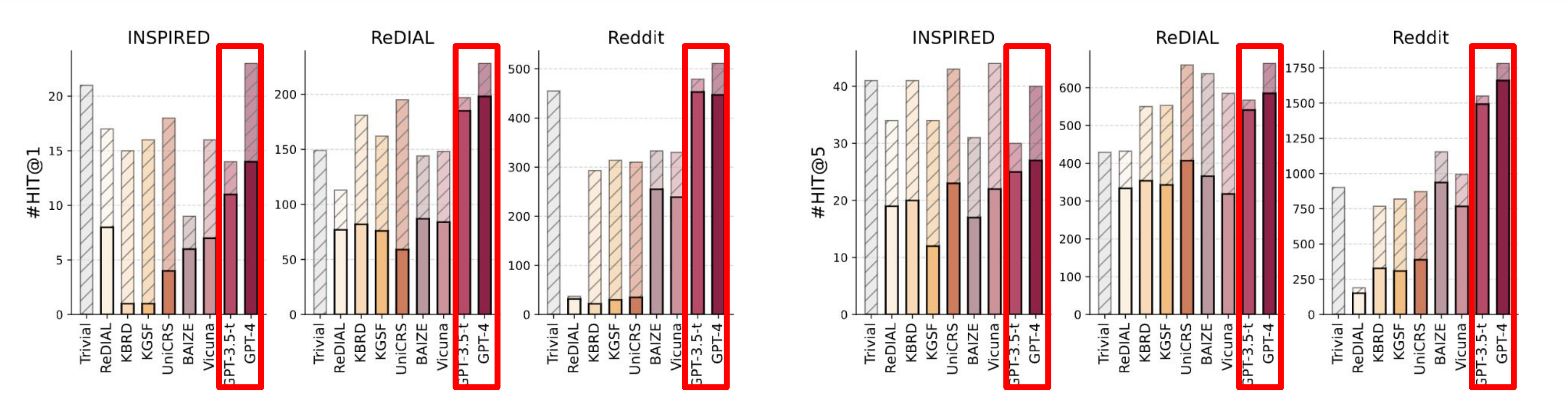


How powerful are LLMs for zero-shot CRS?

He et al. Large Language Models as Zero-Shot Conversational Recommenders. CIKM 2023.

LLM as Conversational Recommender

LLMs as zero-shot CRS

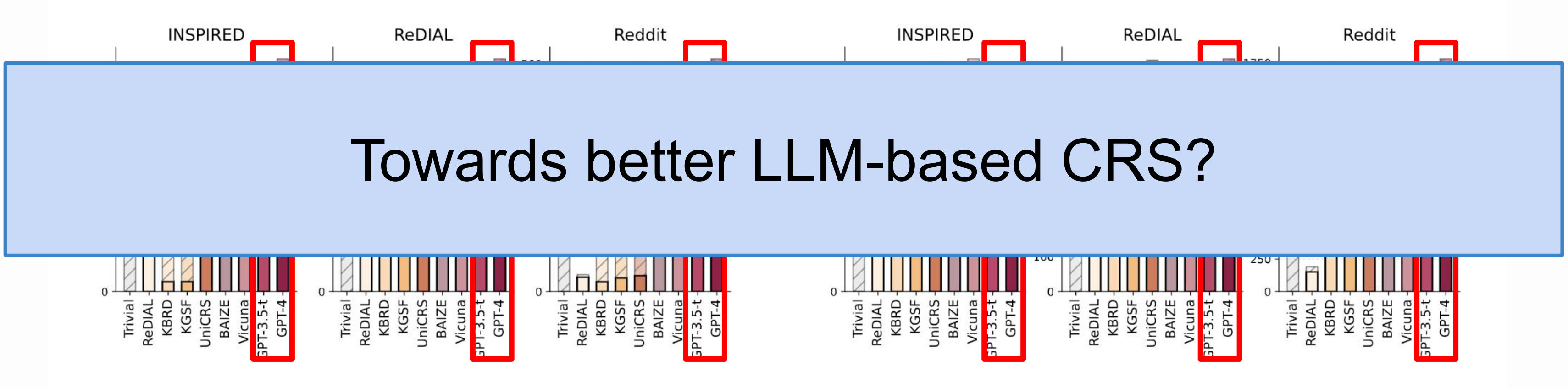


Can surpass traditional CRSs!

He et al. Large Language Models as Zero-Shot Conversational Recommenders. CIKM 2023.

LLM as Conversational Recommender

LLMs as zero-shot CRS

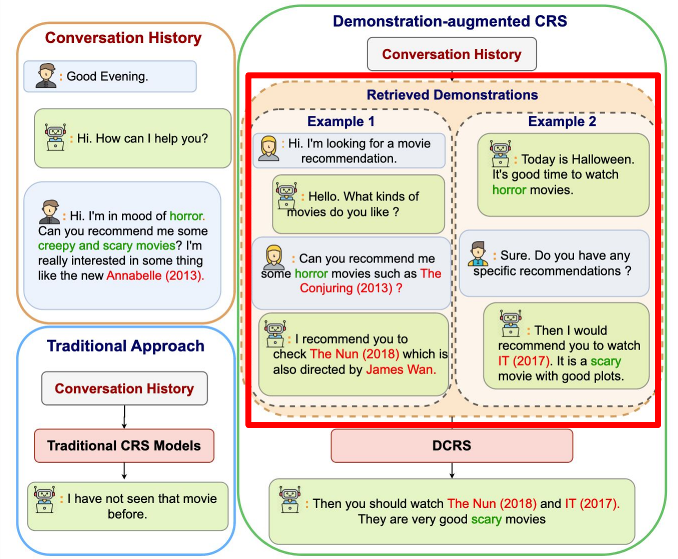


Can surpass traditional CRSs!

He et al. Large Language Models as Zero-Shot Conversational Recommenders. CIKM 2023.

LLM as Conversational Recommender

+ ## Demonstration



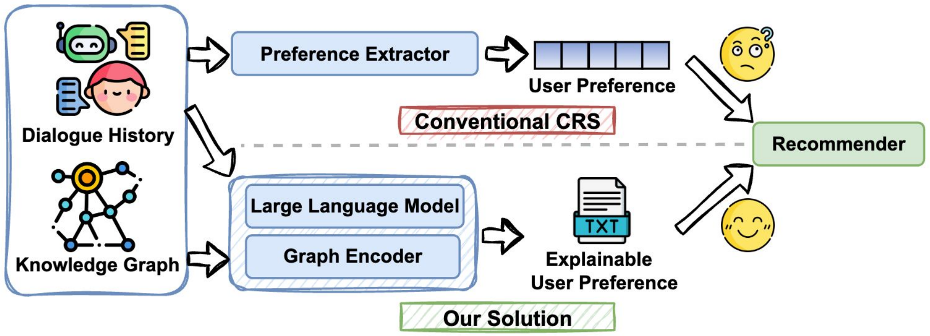
Prompting with previously successful conversation

Relevant conversation history helps!

Dao et al. Broadening the View: Demonstration-augmented Prompt Learning for Conversational Recommendation. SIGIR 2024.

LLM as Conversational Recommender

+ Knowledge graph

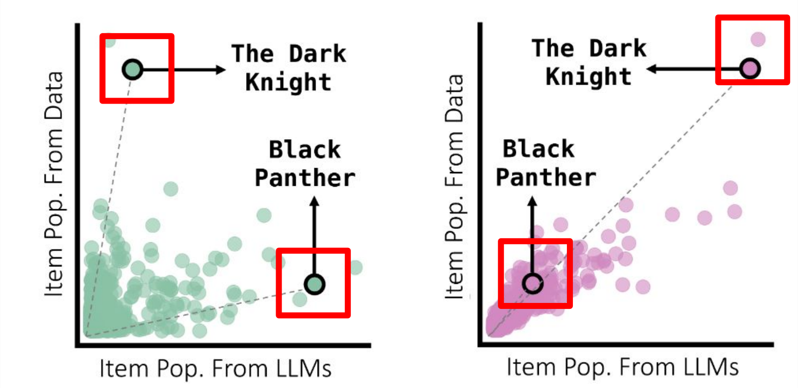


Recommendation-specific knowledge graph helps

Qiu et al. Unveiling User Preferences: A Knowledge Graph and LLM-Driven Approach for Conversational Recommendation. arXiv:2411.14459

LLM as Conversational Recommender

+ ## Collaborative information



(a) Before *RTA* (b) After *RTA*

Collaborative information (e.g., popularity) helps LLMs fit the real distribution in CRS

He et al. Reindex-Then-Adapt: Improving Large Language Models for Conversational Recommendation. WSDM 2025.

LLM as Conversational Recommender

Challenges – Datasets

Public datasets for CRS are limited, due to the scarcity of conversational products and real-world CRS datasets

LLM as Conversational Recommender

Challenges – Evaluation

Traditional metrics like NDCG and BLEU are often insufficient to assess user experience

LLM as Conversational Recommender

Challenges – Product

What is the form of LLM-based CRS products?

ChatBot? Search bar? Independent App?

Part 2: LLM as Conversational Recommender

(1) LLMs show potential in CRS

(2) LLM-based CRS can be improved with:

demonstration, collaborative information …

(3) Challenges in LLM-based CRSs:

dataset, evaluation, and product

Part 3: LLM as User Simulator

User simulators before the era of LLM

RL-based user simulator

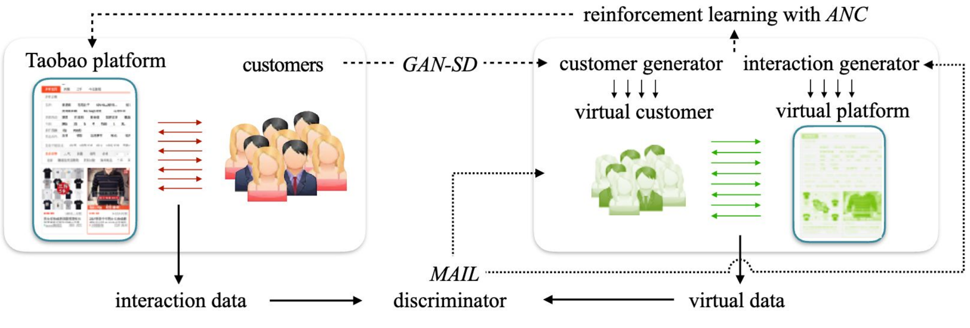


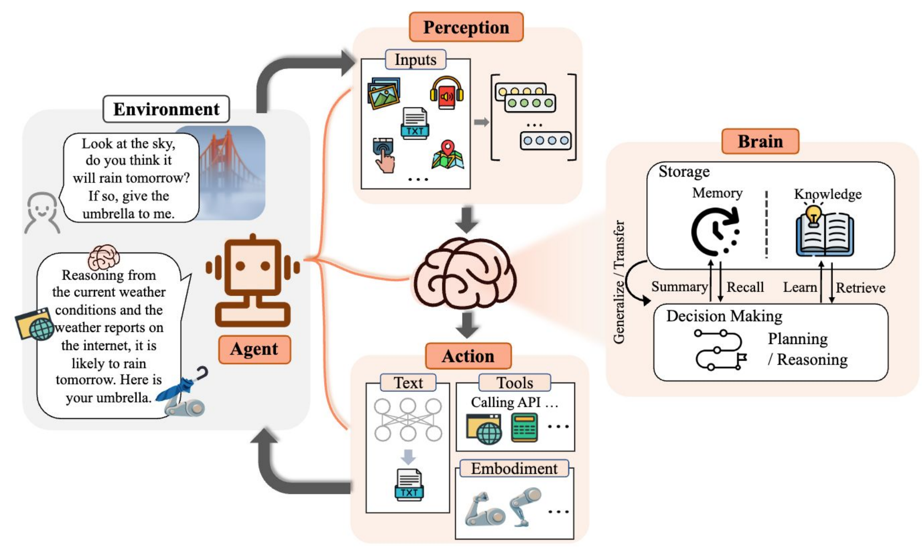
Figure 1: Virtual-Taobao architecture for reinforcement learning.

High sampling cost
Overfitting risks
Training instability
Limited action space

...

LLM as User Simulator

Generative agents



Perception

Planning

Memory

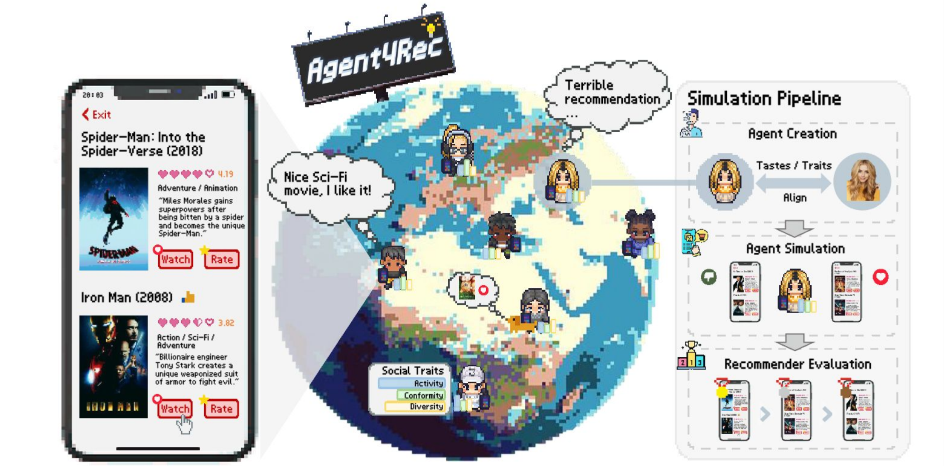
Action

...

Xi et al.The Rise and Potential of Large Language Model Based Agents: A Survey. arXiv: 2309.07864.

LLM as User Simulator

Generative agents for recommendation

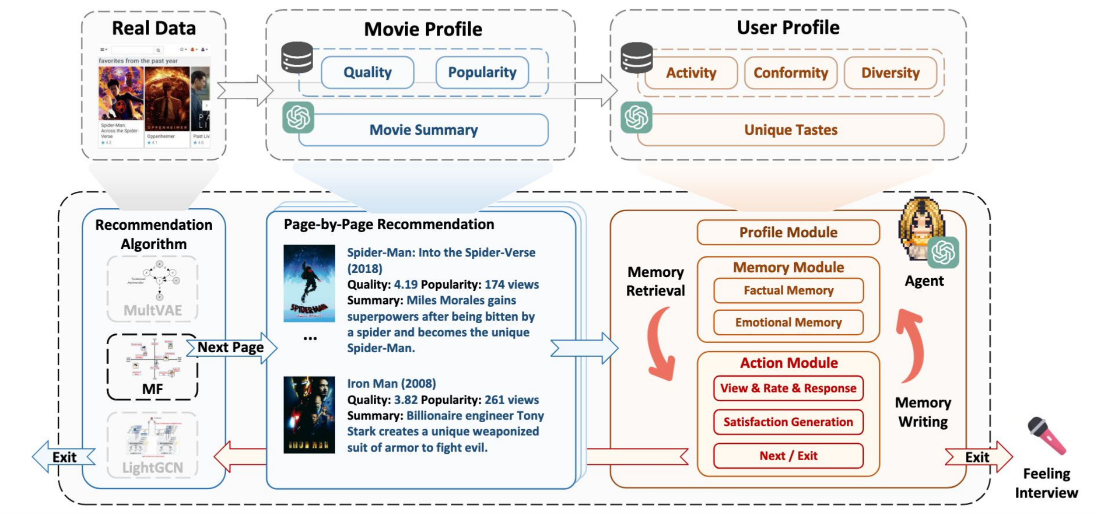


Human-like behavior
Abundant action space
Reduced training cost
…

Zhang et al. On generative agents in recommendation. SIGIR 2024.

LLM as User Simulator

Generative agents for recommendation



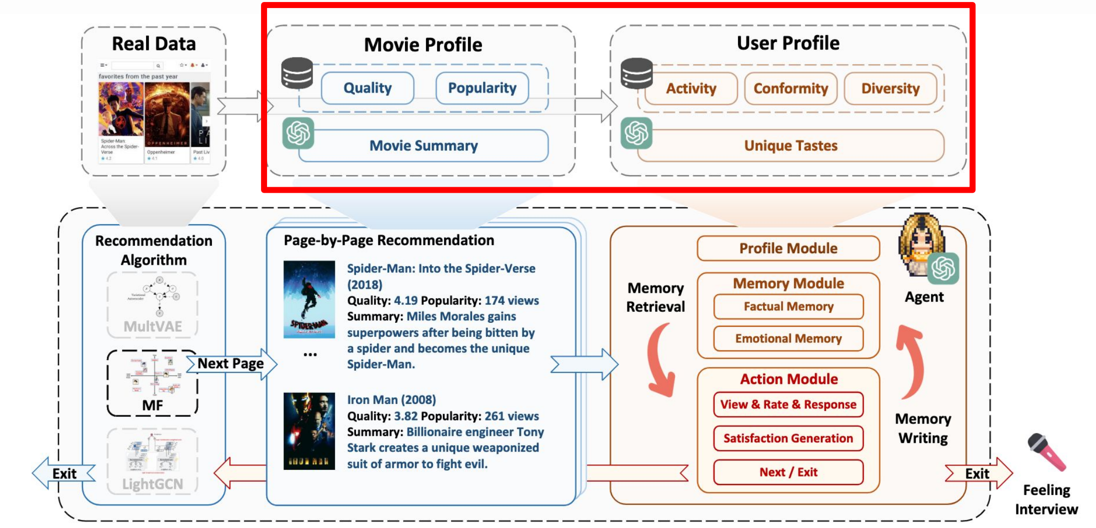
Realworld-like simulation paradigm

- 1000 users
- Page-by-page simulation

Zhang et al. On generative agents in recommendation. SIGIR 2024.

LLM as User Simulator

Generative agents for recommendation



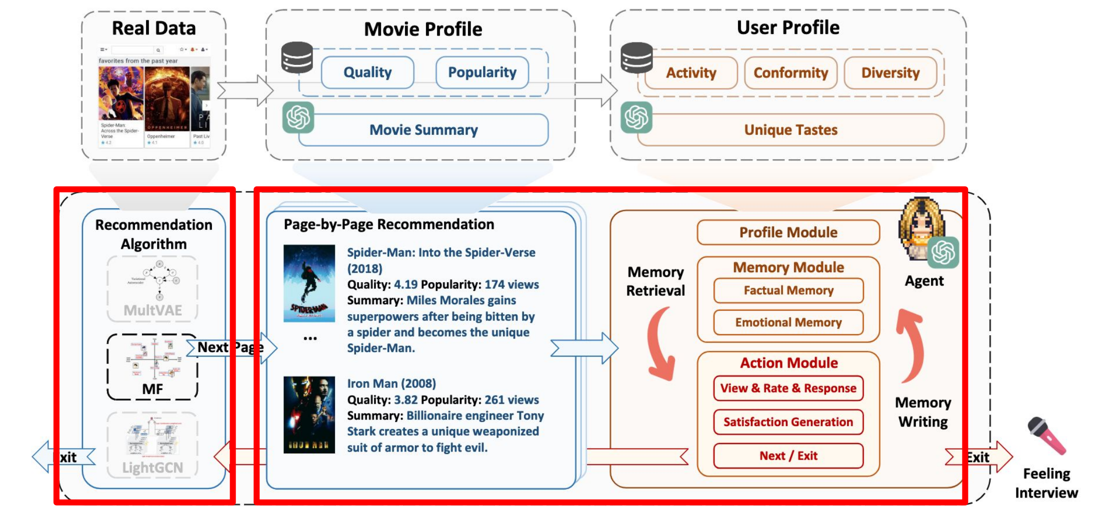
Realworld-like simulation paradigm

- 1000 users
- Page-by-page simulation

Zhang et al. On generative agents in recommendation. SIGIR 2024.

LLM as User Simulator

Generative agents for recommendation



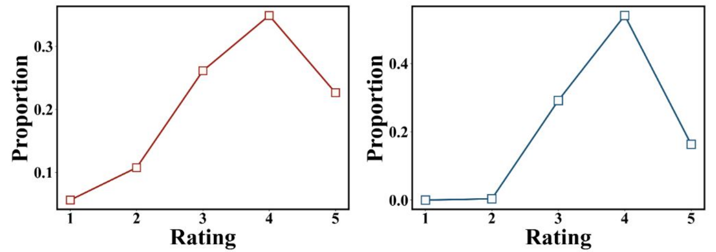
Realworld-like simulation paradigm

- 1000 users
- Page-by-page simulation

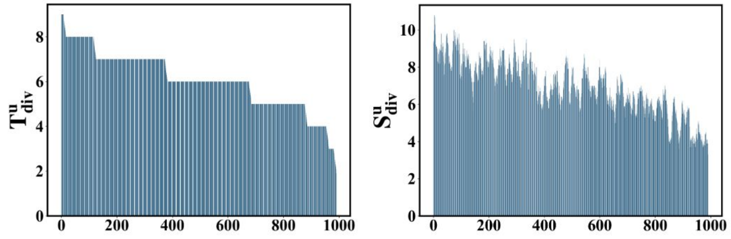
Zhang et al. On generative agents in recommendation. SIGIR 2024.

LLM as User Simulator

Generative agents for recommendation



(a) Distribution on MovieLens (b) Agent-simulated distribution



(a) Ground-truth diversity (b) Simulated diversity

Aligned user preferences
& Recommender evaluation

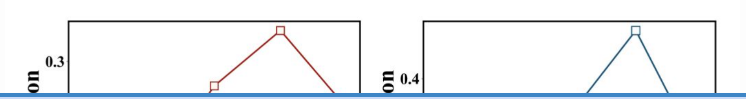
Table 2: Recommendation strategies evaluation.

	$\overline{P}_{view}$	$\overline{N}_{like}$	$\overline{P}_{like}$	$\overline{N}_{exit}$	$\overline{S}_{sat}$
Random	0.312	3.3	0.269	2.99	2.93
Pop	0.398	4.45	0.360	3.01	3.42
MF	0.488	**6.07***	0.462	**3.17***	3.80
MultVAE	0.495	5.69	0.452	3.10	3.75
LightGCN	**0.502***	5.73	**0.465***	3.02	**3.85***

Zhang et al. On generative agents in recommendation. SIGIR 2024.

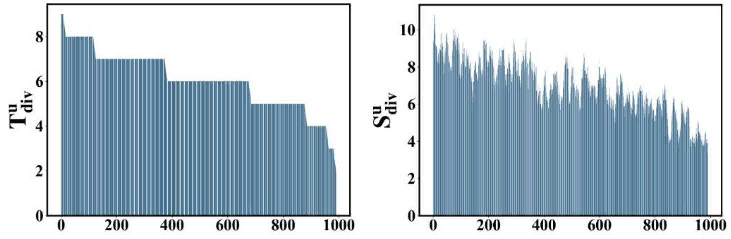
LLM as User Simulator

Generative agents for recommendation



Aligned user preferences

How to make the simulation more faithful?



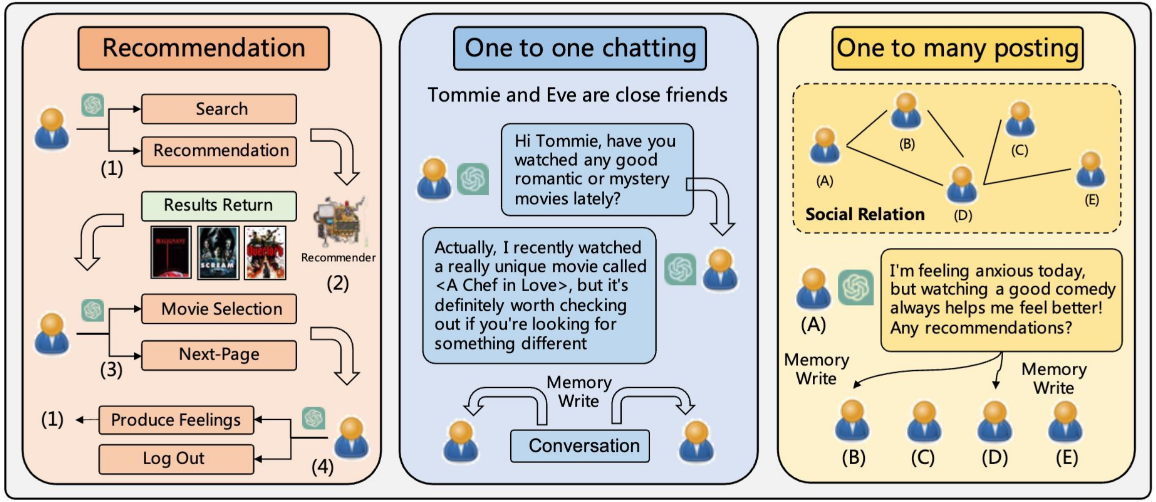
(a) Ground-truth diversity (b) Simulated diversity

	$\overline{P}_{view}$	$\overline{N}_{like}$	$\overline{P}_{like}$	$\overline{N}_{exit}$	$\overline{S}_{sat}$
Random	0.312	3.3	0.269	2.99	2.93
Pop	0.398	4.45	0.360	3.01	3.42
MF	0.488	**6.07***	0.462	**3.17***	3.80
MultVAE	0.495	5.69	0.452	3.10	3.75
LightGCN	**0.502***	5.73	**0.465***	3.02	**3.85***

Zhang et al. On generative agents in recommendation. SIGIR 2024.

LLM as User Simulator

+ ## Social behaviors



Recommendation
Chat
Networking

Wang et al. When Large Language Model based Agent Meets User Behavior Analysis: A Novel User Simulation Paradigm. TOIS 2025.

LLM as User Simulator

+ Multiturn conversation

User ID: A1EXXXXXDUE6B0

Historical Interactions: ["Robin Williams: Live On Broadway", "**Mission Impossible on VHS**", "**Solaris**", "**Elysium**", "Wall Street", "Mystic River", ...]

Conversation 1:

[User] Hi, I recently watched **Mission Impossible on VHS** and it was a fantastic high-tech spy movie! Tom Cruise gave …

[Agent] That's great to hear! I have a movie recommendation for you based on your likes. How about checking out World War Z?

[User] Oh, I'm definitely interested. Can you tell me more about the movie?

[Agent] Certainly! World War Z is a good zombie war flick…

Conversation 2:

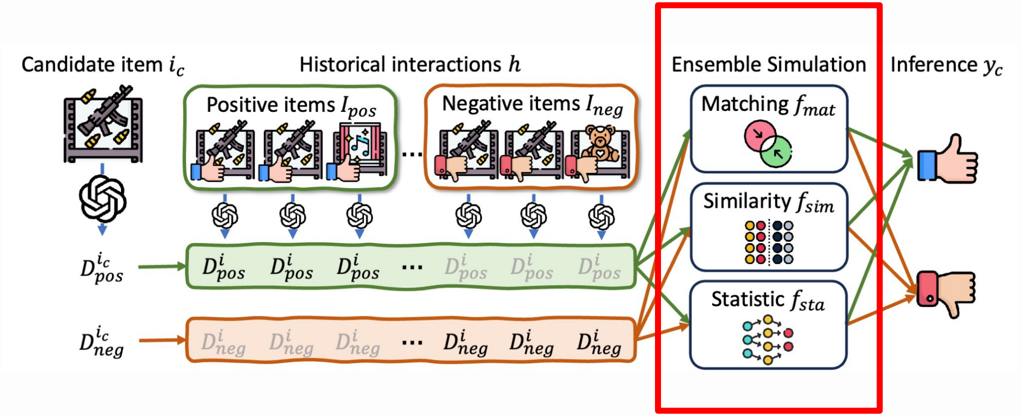
[User] Hi, I recently watched the movie "**Solaris**" and I have to say, I didn't enjoy it at all. It felt overly melodramatic and lacked substance…
[Agent] I understand why you didn't like "Solaris". I can recommend a movie you might enjoy instead. How about "**Elysium**"? It's a bilingual film that…
[User] Oh, I've actually already watched "Elysium" and it was better than I expected, but still not great.…
[Agent] How about giving "Wrecking Crew" a try? It's another movie you might like based on your previous preferences…
[User] Sure, that sounds interesting…

Simulating users in the conversational scenarios

Liang et al. LLM-REDIAL: A Large-Scale Dataset for Conversational Recommender Systems Created from User Behaviors with LLMs. ACL 2024.

LLM as User Simulator

+ ## Multi-facet simulation objective



Category matching
Fine-grained similarity
Statistic information

LLM as User Simulator

+ Multi-facet simulation objective

Dataset	Metric	PPO	TRPO	A2C	DQN
Yelp	**A. Rwd↑**	9.97	13.45	24.15	**27.56**
	T. Rwd↑	141.57	157.42	267.60	**330.98**
	Liking%↑	34.59	40.07	48.35	**49.43**
Amazon Music	**A. Rwd↑**	10.49	11.31	13.45	**16.70**
	T. Rwd↑	129.03	140.15	141.03	**181.42**
	Liking%↑	29.30	32.46	29.54	**33.18**
Amazon Games	**A. Rwd↑**	18.72	21.35	**27.56**	26.43
	T. Rwd↑	208.43	242.26	**317.56**	269.02
	Liking%↑	33.15	37.64	**43.52**	40.73
Amazon Movie	**A. Rwd↑**	29.42	27.47	31.72	**38.60**
	T. Rwd↑	310.69	301.40	354.34	**416.18**
	Liking%↑	38.59	36.70	42.37	**44.50**
Anime	**A. Rwd↑**	14.12	14.58	**21.50**	18.03
	T. Rwd↑	155.74	163.44	**242.95**	201.94
	Liking%↑	25.46	24.27	**31.52**	30.67

Reliable environment for RL-based recommenders

Zhang et al. LLM-Powered User Simulator for Recommender System. AAAI 2025.

Part 3: LLM as User Simulator

(1) RL-based simulators are limited in action space, action space, and training instability

(2) LLMs open up a new paradigm for simulating users

(3) They can give feedback for RL-based recommenders

(4) Challenges:
scaling, training, industry deployment